

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 10:51:49 ; Search time 18 Seconds

(without alignments)
1863.940 Million cell updates/sec

Title: US-09-654-652a-3

Perfect score: 1824

Sequence: 1 MNIKKTAVKSAALVAANAAAA.....AKGAKVNPNGKRRVNFEEH 349

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1824	100.0	349	2 A44507	licheninase (EC 3. endo-beta-1,3-1,4
2	268.5	14.7	851	2 H84053	licheninase (EC 3. licheninase (EC 3.
3	265.5	14.6	252	2 A48378	licheninase (EC 3. licheninase (EC 3.
4	248	13.6	242	1 LKBS	licheninase (EC 3. licheninase (EC 3.
5	247.5	13.6	276	2 I40453	licheninase (EC 3. licheninase (EC 3.
6	242.5	13.3	239	1 A29091	licheninase (EC 3. xylanase, beta(1,3
7	242.5	13.3	802	1 A36910	licheninase (EC 3. licheninase (EC 3.
8	237	13.0	334	1 S23498	licheninase (EC 3. licheninase (EC 3.
9	233.5	12.8	341	1 S15388	licheninase (EC 3. licheninase (EC 3.
10	230	12.6	507	2 S64507	Probable membrane licheninase (EC 3.
11	228	12.5	237	1 S11927	licheninase (EC 3. licheninase (EC 3.
12	227.5	12.5	238	1 S19012	licheninase (EC 3. endo-1,3(4)-beta-g
13	224.5	12.3	242	2 JS0611	endo-1,3(4)-beta-g endo-1,3-1,4-beta-9
14	221	12.1	246	2 D97245	endo-1,3-1,4-beta-9 endo-1,3-1,4-beta-
15	206	11.3	263	2 AB3055	endo-beta-1,3-1,4-endo-beta-1,3-1,4-
16	206	11.3	289	2 C98231	hybrid-endo-beta-1
17	197	10.8	269	2 H95976	UTR2 protein - yea
18	190.5	10.4	302	2 G84053	exoK protein - Rhl
19	189.5	10.4	467	2 S30839	laminaarase - The
20	180.5	9.9	268	2 S34804	hypothetical prote
21	173	9.5	282	2 T02354	xyloglucan endo-1,
22	173	9.5	642	2 B72428	xyloglucan endo-1,
23	172.5	9.5	277	2 B85354	xyloglucan endo-1,
24	170.5	9.3	282	2 A85354	xyloglucan endo-1,
25	164	9.0	286	2 S71225	xyloglucan endo-1,
26	159.5	8.7	287	2 T04236	xyloglucan endo-1,
27	157.5	8.6	269	2 S61555	xyloglucan endo-1,
28	154.5	8.5	277	2 S71222	xyloglucan endo-1,
29	149	8.2	289	2 T06166	xyloglucan endotra

30	146.5	8.0	284	2 T52097	xyloglucan endo-1,
31	145.5	8.0	310	2 A86239	protein T10024.17
32	145	7.9	305	2 G84568	probable xylogluca
33	140.5	7.7	280	2 T02090	xyloglucan endo-1,
34	139.5	7.6	286	2 T06202	xyloglucan endo-1,
35	138	7.6	422	2 S48564	probable membrane
36	137	7.5	292	2 T06201	xyloglucan endo-1,
37	133.5	7.3	286	2 S48201	licheninase (EC 3.
38	132	7.2	295	2 T10210	xyloglucan endo-1,
39	131.5	7.2	301	2 C87296	beta-glucanase [lm
40	129.5	7.1	1324	2 T18265	endo-1,3(4)-beta-g
41	128.5	7.0	277	2 JE0156	end-xyloglucan tra
42	128.5	7.0	292	2 T04514	xyloglucan endo-1,
43	128	7.0	283	2 T07678	xyloglucan endo-1,
44	128	7.0	846	2 AD2672	conserved hypochet
45	128	7.0	848	2 B97454	hypothetical prote

ALIGNMENTS

RESULT 1

A44507

licheninase (EC 3.2.1.73) - Fibrobacter succinogenes

C:Species: Fibrobacter succinogenes

C>Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 08-Oct-1999

C:Accession: A44507

R:Teather, R.M.; Erile, J.D.

J. Bacteriol. 172, 3837-3841, 1990

A:Title: DNA sequence of a Fibrobacter succinogenes mixed-linkage beta-glucanase (1,3

A:Reference number: A44507; MUID:90299807; PMID:2193918

A:Accession: A44507

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-349 <TEA>

A:Cross-references: EMBL:M33676; NID:q148575; PID:AAA24896.1; PID:q148576

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match	100.0%	Score 1824;	DB 2;	Length 349;
Best Local Similarity	100.0%	Pred. No. 1.5e-122;		
Matches	349;	Conservative	0;	Mismatches 0;
		Indels	0;	Gaps 0;
QY	1	MNKKKTAVKSAALVAANAAATTTNWSAKDSGAEIYLEVOYGFARMKMAASGTVS	60	
DB	1	MNKKKTAVKSAALVAANAAATTTNWSAKDSGAEIYLEVOYGFARMKMAASGTVS	60	
QY	61	SMFLYQNGSEIADGRPMVEVDIEVLGRNPGSFQSNITTGAKAGAKTSEKHNAAVSPADQA	120	
DB	61	SMFLYQNGSEIADGRPMVEVDIEVLGRNPGSFQSNITTGAKAGAKTSEKHNAAVSPADQA	120	
QY	121	FHTYGLMTIRNYVWYTDGQEVKTEGGQVSNLTGTGGLRPNLMSSSAAMVGOFDPSKL	180	
DB	121	FHTYGLMTIRNYVWYTDGQEVKTEGGQVSNLTGTGGLRPNLMSSSAAMVGOFDPSKL	180	
QY	181	PLFOFINWVYKKTTPQGGEGSFTLDMTDPNFTPGSRMGKDMTFDGNRPVLDLTKNI	240	
DB	181	PLFOFINWVYKKTTPQGGEGSFTLDMTDPNFTPGSRMGKDMTFDGNRPVLDLTKNI	240	
QY	241	YSRDGMILIALTRKGSFNGQVPRDDEPAPOSSSSAPASSSSVPASSSSVPAASSSAFV	300	
DB	241	YSRDGMILIALTRKGSFNGQVPRDDEPAPOSSSSAPASSSSVPASSSSVPAASSSAFV	300	
QY	301	PPSSSSATIRIHGKRTTPPAVAKERHNLVNAKAKVNPNGKRRVNFEEH 349		
DB	301	PPSSSSATIRIHGKRTTPPAVAKERHNLVNAKAKVNPNGKRRVNFEEH 349		

RESULT 2

endo-beta-1,3-1,4 glucanase (licheninase) bsls [imported] - Bacillus halodurans (str

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: H84053

F:29-242/Product: licheninase #status predicted <MAT>
F:29/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) (partial) #status

Query Match 13.6%; Score 248; DB 1; Length 242;
Best Local Similarity 36.0%; Pred. No. 1.5e-10;
Matches 63; Conservative 25; Mismatches 75; Indels 12; Gaps 7;
QY 20 ALTTNYSAKDFSGAELTYLEEVQYGFEMRMKMAASGVSSMFLYONGSEIADGRPWVE 79
DB 78 ALTSPAYNKK-FDCGERSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE--DGPWDE 133
QY 80 VDIEVLGKNPDSFQSNITITGKAGAKTSEKHNAVSPADQAEHTYTGLEWTPNYVRYMTVDGQ 139
DB 134 IDIEFLGKDTTKYQFNYVTGAGC---NHKEIVDLGSDANAHYTAFFDMQPSIKRYVDS 190
QY 140 QEVKRTGGGVSNLTGTQGLRFNLMSSESA-AWVGQFDESKLPDLFOFINMKVYK 193
DB 191 Q-LKHTATNQIDPTTGGK--IMNMLMNGTGVDEWLGSYNGVN-PLYAHYDMVRYTK 241

RESULT 5

140453
Licheninase (EC 3.2.1.73) - Bacillus sp.
C:Species: Bacillus sp.
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 22-Jun-1999
C:Accession: I40453; S32688
R:Taberero, C.; Coll, P.M.; Fernandez-Abalos, J.M.; Perez, P.; Santamaria, R.I.
Appl. Environ. Microbiol. 60, 1213-1220, 1994
A:Title: Cloning and DNA sequencing of bgal, a gene encoding an endo-beta-1,3-1,4-glucan
A:Reference number: I40453; MUID:94288605; PMID:7517127
A:Accession: I40453
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-276 <RES>
A:Cross-references: EMBL:212151; NID:g296931; PIDN:CAA78135.1; PID:g296932
A:Gene: bgal
C:Superfamily: licheninase
C:Keywords: glycosidase; hydrolase

Query Match 13.6%; Score 247.5; DB 2; Length 276;
Best Local Similarity 36.1%; Pred. No. 2e-10;
Matches 61; Conservative 20; Mismatches 75; Indels 13; Gaps 6;
QY 30 FSGAELTYLEEVQYGFEMRMKMAASGVSSMFLYONGSEIADGRPWVEVDIYLGKRP 89
DB 96 YKAGELRTNDFYHGYGLYEVRMKPAKNTGIVSSFFTY-TGPWDENDPWEIDIEFLGKDT 154
QY 90 GSFQSNITITGKAGAKTSEKHNAVSPADQAEHTYTGLEWTPNYVRYMTVDGQEVKRTGGQ 149
DB 155 TKIQFNYVTGAGC---NEHYHDLGFDADDFYTAFFMRPESIRKRFVNGELVHTA---- 207
QY 150 VSNLTGT-QGLRFNLMSSESA-AWVGQFDESKLPDLFOFINMKVYK 196
DB 208 TENIPQTPQKIMNMMPGIGVDMGTGRFNGEDPPTVYQDWV---KYTP 253

RESULT 6

A29091
Licheninase (EC 3.2.1.73) beta - Bacillus amyloliquefaciens
N:Alternate names: 1,3-1,4-beta-D-glucan 4-glucanohydrolase; beta-glucanase; lichenase
C:Species: Bacillus amyloliquefaciens
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
R:Jofemeister, J.; Kurtz, A.; Borris, R.; Knowles, J.
Gene 49, 177-187, 1986
A:Title: The beta-glucanase gene from Bacillus amyloliquefaciens shows extensive homolog
A:Reference number: A91564; MUID:87192007; PMID:3106158
A:Accession: A29091
A:Molecule type: DNA
A:Residues: 1-239 <HOF>
A:Cross-references: GB:M15674; NID:g143009; PIDN:AAA87323.1; PID:g143010
A:Experimental source: strain BE20/78

C:Genetics:
A:Gene: bgla
C:Superfamily: licheninase
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 13.3%; Score 242.5; DB 1; Length 239;
Best Local Similarity 34.3%; Pred. No. 3.7e-10;
Matches 59; Conservative 27; Mismatches 75; Indels 11; Gaps 6;
QY 23 TNVSAKDFSGAELTYLEEVQYGFEMRMKMAASGVSSMFLYONGSEIADGRPWVEYDI 82
DB 77 TSPSYNKKFDCGERSVQTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE--CTPWEIDDI 133
QY 83 EVLGRKPSFQSNITITGKAGAKTSEKHNAVSPADQAEHTYTGLEWTPNYVRYMTVDGQEV 142
DB 134 EFLGKDTTKYQFNYVTGAGC---NHKEFADLGFDAANAHYTAFFDMQPSIKRYVDSGO-L 189
QY 143 RTEGGQVSNLTGTQGLRFNLMSSESA-AWVGQFDESKLPDLFOFINMKVYK 193
DB 190 KHTATNQIDPTTGGK--IMNMLMNGTGVDEWLGSYNGVN-PLYAHYDMVRYTK 238

RESULT 7

A36910
Xylanase, beta(1,3-1,4)-glucanase - Ruminococcus flavefaciens
C:Species: Ruminococcus flavefaciens
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999
C:Accession: A36910
R:Plint, H.J.; Martin, J.; McPherson, C.A.; Daniel, A.S.; Zhang, J.X.
J. Bacteriol. 175, 2943-2951, 1993
A:Title: A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-glucanase do
A:Reference number: A36910; MUID:93259938; PMID:8491715
A:Accession: A36910
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-802 <FLI>
A:Cross-references: GB:S61204; NID:g385910; PIDN:AA826620.1; PID:g385911
A:Note: sequence extracted from NCBI backbone (NCBIN:131871, NCBI:P:131872)
C:Superfamily: endo-1,4-beta-xylanase homology; Thermotoga xylanase A amino-terminal
F:42-239/Domain: endo-1,4-beta-xylanase homology <XLY>
F:259-401/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

Query Match 13.3%; Score 242.5; DB 2; Length 802;
Best Local Similarity 34.9%; Pred. No. 1.7e-09;
Matches 58; Conservative 21; Mismatches 74; Indels 13; Gaps 6;
QY 30 FSGAELTYLEEVQYGFEMRMKMAASGVSSMFLYONGSEIADGRPWVEVDIYLGKRP 89
DB 634 YSGGEFRTNMFYHGYGLYEVRMKPAKNTGIVSSFFTYTGPS--DNPWDEIDIEFLGKNT 690
QY 90 GSFQSNITITGKAGAKTSEKHNAVSPADQAEHTYTGLEWTPNYVRYMTVDGQEVKRTGGQ 149
DB 691 TOVQFNYVTGAGC---EKLVDLGFDSSEAHYTGFDQPNYIMAYDGRVRYA---- 743
QY 150 VSNLTGT-QGLRFNLMSSESA-AWVGQFDESKLPDLFOFINMKVYK 193
DB 744 TQDIPRTQKIMNMMPGLTVDMLKATN-GRTPLFAHQVYTYK 788

RESULT 8

S23498
Licheninase (EC 3.2.1.73) l1cb precursor - Clostridium thermocellum
N:Alternate names: beta-1,3-1,4-glucanase l1cb; lichenase l1cb
C:Species: Clostridium thermocellum
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S23498; S23137
R:Schlamm, S.; Schwarz, W.H.; Staudenbauer, W.L.
Eur. J. Biochem. 204, 33-39, 1992
A:Title: Structure of the Clostridium thermocellum gene l1cb and the encoded beta-1,3
l cellulases.
A:Reference number: S23498; MUID:92155194; PMID:1740123
A:Accession: S23498
A:Molecule type: DNA

A:Residues: 1-334 <SCH>
 A:Cross-references: EMBL:X63355; NID:g40697; PIDN:CAA44959.1; PID:g40698
 C:Genetics:
 A:Gene: l1cB
 C:Superfamily: licheninase l1cB: Clostridium cellulase repeat homology
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-334/Product: licheninase #status predicted <NMT>
 F:273-296/Domain: Clostridium cellulase repeat homology <CCRI>
 F:308-331/Domain: Clostridium cellulase repeat homology <CCR1>

Query Match 13.08; Score 237; DB 1; Length 334;
 Best Local Similarity 30.98; Pred. No. 1,4e-09;
 Matches 80; Conservative 27; Mismatches 94; Indels 58; Gaps 13;

QY 28 KDFSGAEIYLTLEEVQ-----YGFKEARMKMAASGVVSMFLYONGSEIADRPVYED 81
 DB 82 REVGGSIPYKSGEYKRSFEGYGYEYKMAKNGVSSFFYTTGPS---DNNPWEID 138
 QY 82 IEVLKKNPGSFQSNITGKAGAOCTSEKHAHVSPADQAFHTYGLETPVNVVMTVDGQE 141
 DB 139 IEFLGKDTTKYQFWMYKNGVG---NEYLHNLGFDASQDFHTYGFEMRPDYIDFYDGKK 195
 QY 142 VRKTEGQVSNLGTGQ-LRFNLMSSSA-AMVGQFDESKLPFQFLIMYKYV----- 192
 DB 196 VYR---GTRNIPVTPGKIMMNLMPGIGVDEWLGKRD-GRTPLOAEYEVYKYPNGVPOD 250
 QY 193 -----KYTP-----GQGE-GGSDF-----LDWTDNFDPTDGSRMKG 224
 DB 251 NPTPTPIASTPSPNPLPLKGDVNGDGHVSSDYLKRYLRLVIDRPFVGDQS---VA 307
 QY 225 DWTFDGNRVDLTOKNIYSR 243
 DB 308 DVNRDG-RIDSTDLTMLKR 325

RESULT 9

S15388
 licheninase (EC 3.2.1.73) - Bacillus licheniformis
 N:Alternate names: endo-beta-1,3-1,4-D-glucanase; lichenase
 C:Species: Bacillus licheniformis
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 R:Lloberas, J.; Perez-Pons, J.A.; Querol, E.
 Eur. J. Biochem. 197; 337-343, 1991
 A:Title: Molecular cloning, expression and nucleotide sequence of the endo-beta-1,3-1,4-
 A:Reference number: S15388; MUID:91224124; PMID:2026156
 A:Accession: S15388
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-243 <LLO>
 A:Cross-references: EMBL:X57279; NID:g39558; PIDN:CAA40547.1; PID:g39559
 C:Superfamily: licheninase
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 12.88; Score 233.5; DB 1; Length 243;
 Best Local Similarity 33.78; Pred. No. 1.7e-09;
 Matches 58; Conservative 26; Mismatches 77; Indels 11; Gaps 6;

QY 23 TNVSADFGSAELTYTEEVQKFEARMKMAASGVVSMFLYONGSEIADRPVYED 82
 DB 81 TSPSYNKFQDGENRSVOTYGYGLYVNMKPAKNGVSSFFYTTGPT---DGPWEID 137
 QY 83 EVLGKKNPGSFQSNITGKAGAOCTSEKHAHVSPADQAFHTYGLETPVNVVMTVDGQE 142
 DB 138 EFLGKDTTKYQFWMYKNGVG---NHEKIVNLGFDANSYHTVAFDQNPISIKMYVDGQ-L 193
 QY 143 RKEGQVSNLGTGQGLRFLNLMSSSA-AMVGQFDESKLPFQFLIMYKYV 193
 DB 194 KHTATTOIPQTPGK--IMNLMNGAGVDEWLGKYN-GVTPLSRLSLHMTVRTK 242

RESULT 10

S64507
 Probable membrane protein YGR189c - Yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein G7553
 C:Species: Saccharomyces cerevisiae
 C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 19-Apr-2002
 C:Accession: S64507
 R:Arroyo, J.; Garcia-Gonzalez, M.; Garcia-Saez, M.I.; Sanchez-Perez, M.; Nombela, C.
 Submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64499
 A:Accession: S64507
 A:Molecule type: DNA
 A:Residues: 1-507 <ARR>
 A:Cross-references: EMBL:Z72974; NID:g1323335; PID:e243566; PID:g1323336; GSPDB:GN000
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:CRH1; MIPS:YGR189c
 A:Cross-references: SGD:S0003421
 A:Map position: 7R
 C:Keywords: transmembrane protein
 F:6-22/Domain: transmembrane #status predicted <TM>

Query Match 12.68; Score 230; DB 2; Length 507;
 Best Local Similarity 26.98; Pred. No. 7.5e-09;
 Matches 87; Conservative 57; Mismatches 134; Indels 46; Gaps 13;

QY 6 TAVKSALAVAAAAALTTNVSADFGSAELTYTEEVQKFEARMKMAASGVVSMFLY 65
 DB 70 TDLKHAKEIKYSGDSGLSMTL-AKRYDNPSLKNFIMYGLKLEVLKANGVYSSFLQ 128
 QY 66 QNGSEIADRPVYEDIEVLKKNPGSFQSNITGKAGAOCTSEKHAHVSPADQAFHTY 125
 DB 129 SDLDL-----EIDIEWGSDNTQFQSNFQSKDITTYRGEHVDPTDK-FHNYT 179
 QY 126 LEWTPVNVVMTVDGQEVKTEGQVSNLGTGQ-----LRFNLMSS--ESAA---W 171
 DB 180 LDMAMDKTWYLDGESVR-----VLSNTSSSGYQSPMYLMGIMAGGDPDNNAGTIEW 233
 QY 172 V-GQFDESKLPFQFLIMYKYKYPGQEGGSDFTLDWTDNFDPTDGSRMKGDMTFDG 230
 DB 234 ACGETNVDNAPPTWYIEKYIVDYSTGKKYTGDDSGSM-DSIADGSGTIGRDOAQE- 291
 QY 231 NRVDLTKNIYSRDMGLIALTRKG--QESFGQYPRDDEAPQSSAPAPSS--SSVPA 286
 DB 292 -----DRAVLANGSISSTSSSTVSSASSTVSSSVSSASTVSSSVSTVS 345
 QY 287 SSSVPASSSAFVPPSSSATNA 310
 DB 346 SSSSVSSSSTSS--PSSSTATSS 366

RESULT 11

S11927
 licheninase (EC 3.2.1.73) precursor [validated] - Bacillus macerans

N:Alternate names: endo-beta-1,3-1,4-glucanase; lichenase
 C:Species: Bacillus macerans
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Sep-2000
 C:Accession: S11927
 R:Borriess, R.; Bueltner, K.; Maentzsaalae, P.
 Mol. Gen. Genet. 222, 278-283, 1990

A:Title: Structure of the beta-1,3-1,4-glucanase gene of Bacillus macerans: homologie
 A:Reference number: S11927; MUID:91109712; PMID:2274030
 A:Accession: S11927

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-237 <BOR>
 A:Cross-references: EMBL:X55959; NID:g296715; PIDN:CAA9426.1; PID:g296716
 R:Haub, M.; Heinemann, U.
 Submitted to the Brookhaven Protein Data Bank, December 1994

A:Reference number: A67074; PDB:1MAC
 A:Contents: annotation; X-ray crystallography, 2.3 angstroms, residues 26-37, 'P', 39-2
 A>Note: recombinant form with residues 26-41 from Bacillus amyloliquefaciens expresse
 R:Kietel, T.; Simon, O.; Borriess, R.; Heinemann, U.
 Proc. Natl. Acad. Sci. U.S.A. 90, 5287-5291, 1993

RESULT 13

C;superfamily: Licheninase

RESULT 13

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 10:49:59 ; Search time 11 Seconds
(without alignments)
1315.931 Million cell updates/sec

Title: US-09-654-652a-3

Perfect score: 1824
Sequence: 1 MNIKKTAVKSAALVAAAAA.....AKGAKVNPNGHKRYRVNEH 349

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	1824	100.0	349	1	GUB_FIBSU
2	266.5	14.6	259	1	GUB_BACBR
3	248	13.6	242	1	GUB_BACSU
4	242.5	13.3	239	1	GUB_BACAM
5	242.5	13.3	802	1	XYND_RUMFL
6	237	13.0	334	1	GUB_CLOTH
7	233.5	12.8	243	1	GUB_BACLI
8	230	12.6	507	1	YG46_YEAST
9	228	12.5	237	1	GUB_PAEPA
10	227.5	12.5	238	1	GUB_PAEPO
11	197	10.8	269	1	EXOK_RHIME
12	189.5	10.4	347	1	UTR2_YEAST
13	157.5	8.6	269	1	MER5_ARATH
14	133.5	7.3	286	1	GUB_RHOMR
15	128	7.0	283	1	BRU1_SOYEN
16	119	6.5	465	1	EXSH_RHIME
17	117	6.4	397	1	CGKA_ALMCA
18	115	6.3	682	1	EL3B_BACCI
19	108.5	5.9	989	1	SERA_PLAAG
20	107.5	5.9	719	1	PBPA_STRPN
21	107	5.9	484	1	YODO_YEAST
22	104	5.7	962	1	GUNA_PSEFL
23	103.5	5.7	1580	1	ACCB_HUMAN
24	102.5	5.6	879	1	GUNI_CLOTH
25	102	5.6	465	1	EBLC_RHIME
26	102	5.6	650	1	SCD2_NEUCR
27	101	5.5	691	1	VIVI_MAIZE
28	100.5	5.5	1036	1	NIT2_NEUCR
29	100	5.5	748	1	GUNC_PSEFL
30	99	5.4	344	1	GUNC_TREPA
31	99	5.4	525	1	Y487_TREPA
32	99	5.4	560	1	YDEN_ECOLI
33	98.5	5.4	658	1	GUN3_FIBSU

34	98.5	5.4	1655	1	OMP8_RICCN	O9Kka3 r outer mem
35	98	5.4	1845	1	2236_HUMAN	O9u136 homo sapien
36	97.5	5.3	1385	1	CSAA_BACUD	O94516 bacillus th
37	97.5	5.3	1581	1	ACC8_CRICR	O09427 cricetus cr
38	97.5	5.3	1581	1	ACC8_RAT	O09429 ratius norv
39	97	5.3	1770	1	PMPQ_CHLFR	O84419 chlamydia t
40	96.5	5.3	526	1	CAP_YEAST	P17555 saccharomyc
41	96	5.3	371	1	LIG2_PHACH	P49012 phanerocha
42	95.5	5.2	872	1	IP3L_RAT	P42335 ratius norv
43	95.5	5.2	1289	1	CSAB_BACUD	O45753 bacillus th
44	95	5.2	400	1	GUN5_BACAG	O85465 bacillus ag
45	94.5	5.2	451	1	MYBH_DICDI	P34127 dictyostell

ALIGNMENTS

RESULT 1
GUB_FIBSU
ID GUB_FIBSU STANDARD: PRT: 349 AA.
AC P17989:
DT 01-NOV-1990 (Rel. 15, Created)
DT 01-NOV-1990 (Rel. 15, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Mixed linkage beta-
DE glucanase) (Lichenase).
OS Fibrobacter succinogenes (Bacteroides succinogenes).
OC Bacteria; Fibrobacter/Actinobacteria group; Fibrobacter group;
OC Fibrobacter.
OX NCBI_TaxID=833;
RX [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 28-57.
RC STRAIN-Isolate S85;
RC MEDLINE=90299807; PubMed=2193918;
RA Yeather R.M., Erfle J.D.;
RT "DNA sequence of a fibrobacter succinogenes mixed-linkage
RT beta-glucanase (1,3-1,4-beta-D-glucan 4-glucanohydrolase) gene";
RL J. Bacteriol. 172:3837-3841(1990).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M33676; AAA24896.1; -
CC PIR: A44507; A44507.
CC HSSP: P23904; IAJK.
DR InterPro: IPR000757; Glyco_hydro.16.
DR Pfam: PF00722; Glyco_hydro.16; 1.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KM Hydroxylase; Glycosidase; Signal; Repeat.
FT SIGNAL 1 27
FT CHAIN 28 349
FT ACT_SITE 79 79
FT ACT_SITE 83 83
FT ACT_SITE 83 83
FT DOMAIN 271 307
FT REPEAT 271 277
FT REPEAT 278 284
FT REPEAT 285 291
FT REPEAT 292 298
FT REPEAT 301 307
SQ SEQUENCE 349 AA: 37737 MW: 16DC4F5BDEF578A CAC64;
Query Match 100.0%; Score 1824; DB 1; Length 349.
Best Local Similarity 100.0%; Pred. No. 1; 9e-123;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MNKRTAVKSAALAAAAAALTNNVSAKDFSGAELTYLEEVGYGFEARMKMAASCTVS 60
DB 1 MNKRTAVKSAALAAAAAALTNNVSAKDFSGAELTYLEEVGYGFEARMKMAASCTVS 60
OY 61 SMFLYONGSEIADRPWEVDIEVLGKNPQSGFOSNIITGKGAOKTSEKHAVSPADQA 120
DB 61 SMFLYONGSEIADRPWEVDIEVLGKNPQSGFOSNIITGKGAOKTSEKHAVSPADQA 120
OY 121 FHTGTEPNTVYVNTVDGQEVKRTGEGQVSNLTGTQGLRPNLMSSEAAVNGQFDESKL 180
DB 121 FHTGTEPNTVYVNTVDGQEVKRTGEGQVSNLTGTQGLRPNLMSSEAAVNGQFDESKL 180
OY 181 PLFOFIWVKVKKYKTPGQEGSDFTLDMTDNFTPDGSRMGKGMWTFDGNRDLTDKNI 240
DB 181 PLFOFIWVKVKKYKTPGQEGSDFTLDMTDNFTPDGSRMGKGMWTFDGNRDLTDKNI 240
OY 241 YSRDGMILALTRKGOESFNQVPRDDEPAPQSSSAPASSSSVPASSSSVPASSSAFV 300
DB 241 YSRDGMILALTRKGOESFNQVPRDDEPAPQSSSAPASSSSVPASSSSVPASSSAFV 300
OY 301 PSSSSATNATHGRTTPAVAKHRNLVNAKGVNPNHGRYVNEFH 349
DB 301 PSSSSATNATHGRTTPAVAKHRNLVNAKGVNPNHGRYVNEFH 349

RESULT 2
GUB_BACBR STANDARD: PRT: 259 AA.
ID GUB_BACBR STANDARD: PRT: 259 AA.
AC P37073:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
CN BGLB.
OS Bacillus brevis.
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.
OX NCBI_TaxId=1393;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=93159752; PubMed=7763386;
RA Louw M.E., Reid S.J., Watson T.G.;
RT "Characterization, cloning and sequencing of a thermostable
RT endo-(1,3-1,4) beta-glucanase-encoding gene from an alkalophilic
RT Bacillus brevis."
RL Appl. Microbiol. Biotechnol. 38:507-513(1993).
CC -1- FUNCTION: HYDROLYSES B-GLUCANS CONTAINING MIXED BETA-1,3 AND
CC BETA-1,4 LINKAGES. THIS ENZYME IS THERMOSTABLE: ITS OPTIMAL
CC TEMPERATURE IS NEAR 65-70 DEGREES CELSIUS.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M84339; AAA22265.1; -.
DB HSSP; P23904; IACK.
DB InterPro: IPR000757; Glyco_hydro_16.
DB Pfam: PF00722; Glyco_hydro_16; 1.
DB PRINTS: PR00737; GLYDRLASE16.
DB PROSITE: PS01034; GLYCOSYL-HYDROL_F16; 1.
KM Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 259 BETA-GLUCANASE.

```

```

FT ACT SITE 142 142 NUCLEOPHILE (BY SIMILARITY).
FT ACT SITE 146 146 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 259 AA; 29960 MM; A63C09F281FF5D13 CMC64;
Query Match 14.6%; Score 266.5; DB 1; Length 259;
Best Local Similarity 36.0%; Pred. No. 2,9e-12;
Matches 63; Conservative 24; Mismatches 75; Indels 13; Gaps 6;
OY 26 SAKDFSGAELTYLEEVGYGFEARMKMAASCTVSSMFLYONGSEIADRPWEVDIEVL 85
DB 90 SARNKAGELTNDQFYHGLFEVSMKPAKVEGVTSSEFTY-TGEWMDGDFWDEIDLEFL 148
OY 86 GKNPQSGFOSNIITGKGAOKTSEKHAVSPADQAFHTYGLWTPNTVYVNTVDGQEVKRT 145
DB 149 GKDTTRIQNFNTGVGSG---NEFYDGLGFASSEFNTYAFEMWEDSTITVYVNGEAVHTA 205
OY 146 EGGQVSNLTGT-QGLRPNLMSSESA-AVNGQFDESKLPLFOFIWVKVKKYKTPGO 198
DB 206 ---TENIPQTPQKIMNMLMPGVGDWGTGVFDGDNTPVSYDWMV---RTPLQ 253

RESULT 3
GUB_BACSU STANDARD: PRT: 242 AA.
ID GUB_BACSU STANDARD: PRT: 242 AA.
AC P04957:
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
CN BGLS OR BGL OR LICS OR N15B.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1423;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=C120;
RX MEDLINE=84272222; PubMed=6087283;
RA Murphy N., McConnell D.J., Cantwell B.A.;
RT "The DNA sequence of the gene and genetic control sites for the
RT excreted B. subtilis enzyme beta-glucanase."
RL Nucleic Acids Res. 12:5355-5367(1984).
RN 12
RP SEQUENCE FROM N.A.
RX STRAIN=HL-25;
RA Tezuka H., Yuki T., Yabuchi S.;
RT "Construction of a beta-glucanase hyperproducing Bacillus subtilis
RT using the cloned beta-glucanase gene and a multi-copy plasmid."
RL Agric. Biol. Chem. 53:2335-2339(1989).
RN 13
RP SEQUENCE FROM N.A.
RX STRAIN=168 / BGSCLAI;
RX MEDLINE=97124196; PubMed=8965509;
RA Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA Miwa Y., Fujita Y.;
RT "Sequencing of a 65 kb region of the Bacillus subtilis genome
RT containing the ltc and cel loci, and creation of a 177 kb contig
RT covering the gnt-sacxy region."
RL Microbiology 142:3113-3123(1996).
RN 14
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=95219081; PubMed=7704256;
RA Wolf M., Geczi A., Simon O., Borriass R.;
RT "Genes encoding xylan and beta-glucan hydrolysing enzymes in Bacillus
RT subtilis: characterization, mapping and construction of strains
RT deficient in lichenase, cellulase and xylanase."
RL Microbiology 141:281-290(1995).
RN 15
RP SEQUENCE OF 1-6 FROM N.A.
RX STRAIN=BR151;
RX MEDLINE=96178961; PubMed=8606172;
RA Schmetz K., Stuelke J., Gertz S., Krueger S., Kriegl M., Hecker M.,

```

RA Rak B.;
 RT "lact" a Bacillus subtilis transcriptional antiterminator protein of
 the BglG family";
 RL J. Bacteriol. 178:1971-1979(1996).
 RN [6]
 RP SEQUENCE OF 29-63.
 RA Yuki T., Tezuka H., Yabuuchi S.;
 RT "Purification and some properties of two enzymes from a beta-glucanase
 hyperproducing strain, Bacillus subtilis HL-25";
 RL Agric. Biol. Chem. 53:2341-2346(1989).
 CC In beta-D-glucans containing 1,3- and 1,4-bonds.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X00754; CAA25328.1; -
 DR EMBL: D00518; BAA00405.1; -
 DR EMBL: D83026; BAA11697.1; -
 DR EMBL: Z46862; CAA86922.1; -
 DR EMBL: Z28340; CAA82195.1; -
 DR EMBL: Z99124; CAB15943.1; -
 DR PIR: A22914; LXBS.
 DR PIR: J00110; J00110.
 DR HSSP: P27051; 1GBG.
 DR Sddulstst: B610476; bglS.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR PRINTS: PR00737; GLHYDRLASE16.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 DR KMW Hydrolyase: Glycosidase; Signal; Complete proteome.
 FT SIGNAL 1 28
 FT CHAIN 29 242
 FT MOD_RES 29 29
 FT ACT_SITE 133 133
 FT ACT_SITE 137 137
 FT DISULFID 60 89
 FT VARIANT 24 24
 FT VARIANT 83 83
 FT VARIANT 204 204
 FT SEQUENCE 242 AA; 27268 MW; 45958DEA70F22B29 CRC64;
 Query Match 13.68; Score 248; DB 1; Length 242;
 Best Local Similarity 36.08; Pred. No. 5.6e-11;
 Matches 63; Conservative 25; Mismatches 75; Indels 12; Gaps 7;
 Oy 20 ALTTNYSAKPFGSGAELTYLLEVOYGRKFEARMKMAAAGCTVSSMFLYONGSEIADGRPWE 79
 Db 78 ALTSAPYAK-FDCGENSESVOTYTGILEYVRMKPRAKNTGIYSSFTYGPTE---DGTWDE 133
 Oy 80 VDIEVLCKNGPSFOSNITITGKAGAKTSEKHNHVASPAADAFHTYGLEMTPNVRYTVDS 139
 Db 134 IDIEFLCKDITTKYQFNYYTNGAG---NHEKIYDLGFPDAAAHYHAYARDMQPNSIKWYVDG 190
 Oy 140 QEVKRTKEGGVSNLTGTGGLRFLNWSESA-AMVQGFDESKLPFLQFINNVKYYK 193
 Db 191 Q-LKHRTATNOIPTTPGK--IMMNLNMGTVDEMIGSYNGVN-PLYAHYDWMRYRK 241
 RESULT 4
 GUB_BACAM STANDARD; PRT; 239 AA.
 ID GUB_BACAM
 AC P07980;

DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-JUN-1994 (Rel. 25, Last annotation update)
 DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
 DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
 GN BGLA.
 OS Bacillus amyloliquefaciens.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCCL_TaxID-1390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BE 20/78;
 RX MEDLINE-87192007; PubMed-3106158;
 RA Hofmeister J., Kurtz A., Borries R., Knowles J.;
 RT "The beta-glucanase gene from Bacillus amyloliquefaciens shows
 extensive homology with that of Bacillus subtilis.";
 RL Gene 49:177-187(1986).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
 CC in beta-D-glucans containing 1,3- and 1,4-bonds.
 CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M15674; AAA87323.1; -
 DR PIR: A29091; A29091.
 DR HSSP: P27051; 1GBG.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR PRINTS: PR00737; GLHYDRLASE16.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 DR KMW Hydrolyase: Glycosidase; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 239
 FT ACT_SITE 134 134
 FT DISULFID 57 86
 FT SEQUENCE 239 AA; 26928 MW; A76A64268A7AA0B CRC64;
 Query Match 13.38; Score 242.5; DB 1; Length 239;
 Best Local Similarity 34.38; Pred. No. 1.4e-10;
 Matches 59; Conservative 27; Mismatches 75; Indels 11; Gaps 6;
 Oy 23 TNVSARKPFGSGAELTYLLEVOYGRKFEARMKMAAAGCTVSSMFLYONGSEIADGRPWEVDI 82
 Db 77 TSPSYNKFDCGENSESVOTYTGILEYVRMKPRAKNTGIYSSFTYGPTE---GTPWDEIDI 133
 Oy 83 EVLCKNGPSFOSNITITGKAGAKTSEKHNHVASPAADAFHTYGLEMTPNVRYTVDSQEV 142
 Db 134 EFLGCKDITTKYQFNYYTNGAG---NHEKFPADGFPDAAAHYHAYARDMQPNSIKWYVDGQ-L 169
 Oy 143 RKTGEGVSNLTGTGGLRFLNWSESA-AMVQGFDESKLPFLQFINNVKYYK 193
 Db 190 KHTATTOIPAPGK--IMMNLNMGTVDEMIGSYNGVN-PLYAHYDWMRYRK 238
 RESULT 5
 XYND_RUMFL STANDARD; PRT; 802 AA.
 ID XYND_RUMFL
 AC O53317;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Xylanase/beta-glucanase precursor [includes: Endo-1,4-beta-xylanase
 DE (EC 3.2.1.8) (Xylanase); Endo-beta-1,3-1,4 glucanase (EC 3.2.1.73)
 DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase)].
 GN XYND.

```

OS Ruminococcus flavefaciens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;
OC Ruminococcus.
OX NCBI_TaxID=1265;
RN
RP SEQUENCE FROM N.A.
RX STRAIN=17;
RX MEDLINE=93259938; PubMed=8491715;
RA Eliot H.J., Martin J., McPherson C.A., Daniel A.S., Zhang J.-X.;
RT "A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-
RT glucanase domains, encoded by the xynD gene of Ruminococcus
RT flavefaciens."; 175:2943-2951(1993).
RL J. Bacteriol. 175:2943-2951(1993).
CC -1- FUNCTION: CONTAINS TWO CATALYTIC DOMAINS WITH XYLANASE AND ENDO-
CC BETA-1,3-1,4 GLUCANASE ACTIVITIES.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic
CC linkages in xylans.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- PATHWAY: Xylan degradation.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY
CC G (FAMILY 11 OF GLYCOSYL HYDROLASES).
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO FAMILY 16 OF
CC GLYCOSYL HYDROLASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: S61204; AAB26620.1;
DR HSSP: P23904; IATK.
DR InterPro: IPR003305; CBM_Cenc.
DR InterPro: IPR001137; GH_11.
DR Pfam: PF00457; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR Pfam: PF02018; CBM_4_9; 1.
DR PRINTS: PR00911; GLHYDRASE11.
DR PRINTS: PR00737; GLHYDRASE16.
DR PROSITE: PS00776; GLYCOSYL_HYDROL_F11_1; 1.
DR PROSITE: PS00777; GLYCOSYL_HYDROL_F11_2; 1.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Xylan degradation; Hydrolase; Glycosidase; Signal;
KW Multifunctional enzyme.
FT SIGNAL 1 31
FT CHAIN 32 802
FT DOMAIN 32 244
FT DOMAIN 245 523
FT DOMAIN 524 555
FT DOMAIN 556 802
FT ACT_SITE 124 124
FT ACT_SITE 124 124
FT ACT_SITE 226 226
FT ACT_SITE 684 684
FT DOMAIN 524 529
FT DOMAIN 532 543
FT DOMAIN 546 553
FT SEQUENCE 802 AA; 89091 MW; 28804689647284AF CRC64;
Query Match 13.3%; Score 242.5; DB 1; Length 802;
Best Local Similarity 34.9%; Pred. No. 6e-10;
Matches 58; Conservative 21; Mismatches 74; Indels 13; Gaps 6;

```

```

QY 150 VSNLGTQG-LRENLMSSSA-AWVGFDSEKLPFGFIMWVKYK 193
DB 744 TODIPKTPGKIMNNAWPELTVDWIKAFN-GRTPLAHYQWVYTK 788
RESULT 6
GUB_CLOTM STANDARD: PRT: 334 AA.
AC P29716; P37074;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase) (Laminarinase).
GN LICB OR LAM1.
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1515;
RN
RP SEQUENCE FROM N.A.
RP STRAIN=ATCC 27405 / DSM 1237;
RX MEDLINE=92155194; PubMed=1740123;
RA Schimming S., Schwarz W.H., Staudenbauer W.L.;
RT "Structure of the Clostridium thermocellum gene licB and the encoded
RT beta-1,3-1,4-glucanase. A catalytic region homologous to Bacillus
RT lichenases joined to the reiterated domain of clostridial
RT cellulases.";
RL Eur. J. Biochem. 204:13-19(1992).
RN [12]
RP PRELIMINARY SEQUENCE FROM N.A.
RP STRAIN=F7;
RX MEDLINE=92095946; PubMed=1755832;
RA Zverlov V.V., Lapev D.A., Tishkov V.I., Velikhovskaja G.A.;
RT "Nucleotide sequence of the Clostridium thermocellum laminarinase
RT gene.";
RL Biochem. Biophys. Res. Commun. 181:507-512(1991).
RN [13]
RP REVISIONS.
RA Zverlov V.V.;
RL Submitted (Aug-1993) to the EMBL/GenBank/DBJ databases.
RN [14]
RP CHARACTERIZATION.
RC STRAIN=F7;
RA Zverlov V.V., Velikhovskaja G.A.;
RT "Cloning the Clostridium thermocellum thermostable laminarinase gene
RT in Escherichia coli; the properties of the enzyme thus produced.";
RL Biotechnol. Lett. 12:811-816(1990).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- SUBUNIT: MAY FORM PART OF A MULTIMERIC COMPLEX (CELLULOSE)
CC -1- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS
CC WELL AS IN OTHER C.THERMOCELLUM CELLULOSE ENZYMES. THIS DOMAIN
CC MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X63355; CAA44959.1;
DR EMBL: X58392; CAA41281.1;
DR PIR: S23498; S23498.
DR PIR: JS0611; JS0611.
DR PIR: S18726; S18726.
DR HSSP: P23904; IATK.
DR InterPro: IPR002105; Dockerin_1.
DR InterPro: IPR002048; EF-hand.

```

```

DR InterPro: IPR000757; glyco_hydro_16.
DR Pfam: PF00404; Dockerin_1; 2.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR PRINTS: PR00737; GLYDRLASE16
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_2.
DR PROSITE: PS00448; CLOS_CELLULOSE_RPT; 2.
DR Hydrolyase; Glycosidase; Signal; Repeat.
KW SIGNAL
FT CHAIN 1 27
FT ACT_SITE 136 334 BETA-GLUCANASE.
FT ACT_SITE 140 140 NUCLEOPHILE (BY SIMILARITY).
FT DOMAIN 252 269 PROTON DONOR (BY SIMILARITY).
FT DOMAIN 273 331 PRO/THR-RICH (LINKER).
FT REPEAT 273 296 2 X 24 AA APPROXIMATE REPEATS.
FT REPEAT 308 331 1.
FT REPEAT 304 334 2.
FT CONFLICT 304 334 OSVADVNRDGRIDSDTLMLKRYLRAIPSL ->
SQ SEQUENCE 334 AA: 37897 MW: 0837564E9726F281 CRC64;
Query Match 13.0%; Score 237; DB 1; Length 334;
Best Local Similarity 30.9%; Pred. No. 5.1e-10;
Matches 80; Conservative 27; Mismatches 94; Indels 58; Gaps 13;
OY 28 KDFSGAELTYLEVO-----YGFEEKMAAAGTWSMFLYONGSEIADGRPEVVD 81
DB 82 REYGSYPRKSGEYRTKSFGEYGYEVKMAKNGIYSFFTYGPS---DNNPWDEID 138
OY 82 IEVIGKNGSFQSNITIGKAGAKTSEKHAHVPADDAFHRTYGLEMPNRYRWTVGCOE 141
DB 139 IEFGKDTTKYQFMWYKNGVGG---NEYLHNLGFDASQDFHTYGFEMRPDYIDFYDGKK 195
OY 142 VRKTEGGVSNLTGNG-LRENLMSSEA-AWVGQFDSKPLPQFIIMVYV----- 192
DB 196 VYR---GTRIPPTPKIMANLMPGIGVDMLGRYD-GRIPLQAEYVYVYTPNGVQD 250
OY 193 -----KYTP-----GQGE-GGSDFT-----LDWTNDFDFDGSRMKG 224
DB 251 NPTPTPTAPSTPTNPPLPLKGDVNGGCHVNSDYSLEKRYLARIIDFPVGDQS---VA 307
OY 225 DWTFDGNRVDLITDKNIYSR 243
DB 308 DVNRDG-RIDSTDLTMLKR 325
RESULT 7
GUB_BACLI STANDARD: PRT: 243 AA.
ID GUB_BACLI STANDARD: PRT: 243 AA.
AC P27051;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
GN Bgl.
OS Bacillus licheniformis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_TaxID=1402;
CX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91224124; PubMed=2026156;
RA Lloberas J., Perez-Pons J.A., Querol E.;
RL "Molecular cloning, expression and nucleotide sequence of the
RT endo-beta-1,3-1,4-D-glucanase gene from Bacillus licheniformis.
RL Predictive structural analyses of the encoded polypeptide.";
RL Eur. J. Biochem. 197;337-343(1991).
RN [2]
RP REVISIONS.
RA Querol E.;
RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
RX MUTAGENESIS.
RX MEDLINE=92362869; PubMed=1354172;

```

```

RA Planas A., Juncosa M., Lloberas J., Querol E.;
RT "Essential catalytic role of Glu34 in endo-beta-1,3-1,4-D-glucan 4-
RT glucanohydrolase from B. licheniformis as determined by site-directed
RT mutagenesis.";
RL FEBS Lett. 308:141-145(1992).
RN [4]
RP MUTAGENESIS.
RX MEDLINE=94237863; PubMed=8182059;
RA Juncosa M., Pons J., Dot T., Querol E., Planas A.;
RT "Identification of active site carboxylic residues in Bacillus
RT licheniformis 1,3-1,4-beta-D-glucan 4-glucanohydrolase by
RL site-directed mutagenesis.";
RN J. Biol. Chem. 269:14530-14535(1994).
RP [5]
RX X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RA MEDLINE=96063718; PubMed=7589539;
RT Hahn M., Pons J., Planas A., Querol E., Heinemann U.;
RT "Crystal structure of Bacillus licheniformis 1,3-1,4-beta-D-glucan 4-
RL glucanohydrolase at 1.8-A resolution.";
RL FEBS Lett. 374:221-224(1995).
CC -I- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -I- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -I- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: X57279; CAA04547.1; -.
DR PIR: S15388; S15388.
DR PDB: 1GBG; 07-DEC-95.
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR PRINTS: PR00737; GLYDRLASE16.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Hydrolyase; Glycosidase; Signal; 3D-structure.
FT SIGNAL 1 27
FT CHAIN 28 243
FT ACT_SITE 134 134 BETA-GLUCANASE.
FT ACT_SITE 138 138 NUCLEOPHILE.
FT DISULFID 61 90 PROTON DONOR.
FT MUTAGEN 51 51 D->N: 30% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 51 51 D->N: 85% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 92 92 E->Q: 3% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 92 92 E->Q: 50% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 105 105 D->N: 15% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 133 133 D->N: 0.2% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 134 134 D->N: 0.5% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 136 136 E->Q: COMPLETE LOSS OF ACTIVITY.
FT MUTAGEN 138 138 D->N: 65% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 143 143 D->N: 15% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 160 160 E->Q: 15% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 168 168 D->N: 60% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 179 179 D->N: 80% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 190 190 D->N: 70% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 219 219 D->N: NO CHANGE IN ACTIVITY.
SQ SEQUENCE 243 AA: 27435 MW: 651188D9AAD609A5 CRC64;
Query Match 12.8%; Score 233.5; DB 1; Length 243;
Best Local Similarity 33.7%; Pred. No. 6.1e-10;
Matches 58; Conservative 26; Mismatches 77; Indels 11; Gaps 6;
OY 23 TNVSAKDFSGAELTYLEVOYGFEEKMAAAGTWSMFLYONGSEIADGRPEVVDI 82
DB 81 TSPSTNKKDCGSENSVQYTYGLIEVNKKPAKNWGIYSFFTYGPT---DGTWDELDI 137
OY 83 EVLGNPSSFQSNITIGKAGAKTSEKHAHVPADDAFHRTYGLEMPNRYRWTVGCOE 142

```


Db 138 EFLKDTTKVQFNYYTNGVC---NHEKIVNLGPDANSYHTYAFDQWPNISKWYDQ-1.193
Oy 143 RKTEGGOVSNLTGTGOLRFNLMSSESA-AMYGOFDESKLPLFOFINWKKVK 193
Db 194 KHTATTOIPOPGRK-IMMNLNMGAVDEWLGSTN-GVTPLSRSLHWRYTK 242

RESULT 8
YG46_YEAST STANDARD: PRT: 507 AA.
ID YG46_YEAST STRAND: PRT: 507 AA.
AC P53301;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 52.8 kDa protein in BUB1-HIP1 intergenic region.
GN YG4189C OR G7553.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RX MEDLINE=97279231; Pubmed=9133739;
RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,
R Nombela C.;
RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
RL of Saccharomyces cerevisiae chromosome VII.";
CC Yeast 13:357-363(1997).
CC -1- SIMILARITY: SOME, TO YEAST UTR2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z72974; CAA97215.1; -;
DR EMBL: X99074; CAA67525.1; -;
DR HSSP: P23904; IAKK.
DR SGD: S0003421; CRH1.
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
KW Hypothetical protein.
FT DOMAIN 63 66 POLY-SER.
FT DOMAIN 301 310 POLY-SER.
FT DOMAIN 345 357 POLY-SER.
FT DOMAIN 387 391 POLY-SER.
FT DOMAIN 467 470 POLY-SER.
SQ SEQUENCE 507 AA; 52757 MW; 7D7B61F57AEA942C CRC64;

Query Match 12.6%; Score 230; DB 1; Length 507;
Best Local Similarity 26.9%; Pred. NO. 2.7e-09;
Matches 87; Conservative 57; Mismatches 134; Indels 46; Gaps 13;

Oy 6 TAVSALAVAAALTTNVSADKSCAELTYLEEVQGFPEARKMAAGCTVSSMPLY 65
Db 70 TDLKHAGIKYSGSLMTL-AKRYDNPSLSKSNFYIMGKLEVLKANGCIYSSFTLQ 128
Oy 66 QNGSEIADGRPWVEVDIEVLKKNPQSNIITGKAGAKTSEKHAAVSPADQAFTYTG 125
Db 129 SDDLD-----RIDIEWGCDMTQFQSNFFSKGDTTYYDRGEFHGVDTPTDK-FHNYT 179
Oy 126 LEWPNVYKRVTVDOQVARKTEGGOVSNLTGTG-----LRFNLMSS-----ESAA-----W 171
Db 180 LDMAMDKTWTYLLDESVR-----VLSNTSSEGYPOSPMYLMGIMWAGDPDNDAGTIEW 233
Oy 172 V-GGFDESKLPLFOFINWKKVKYKTPGCGEGSDFTLDWNTNPTPDGSRMGKGDWPTDG 230
Db 234 AGGETNTNDAPFTYITEKIVTYTDYSTGKKYTYGDSGSGW-ESIEADGGSTIGRDQAOE- 291

Oy 221 NRVDLTQKNIXSRDGMILALTRKG--QSEFNGOVPRDDEPAPOSSSAPASS--SSVPA 286
Db 292 -----DEAVLANGSGISSSTSSSTYSSASSTYSSSVSTSSASSTYSSSVSTSS 345
Oy 287 SSSVSPASSSAPVPPSSSATNA 310
Db 346 SSSSVSSSSSTSS--PSSSTATSS 366

RESULT 9
GUB_PAEMA STANDARD: PRT: 237 AA.
ID GUB_PAEMA STRAND: PRT: 237 AA.
AC P23904;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE Paenibacillus macerans (Bacillus macerans) (Lichenase).
OS Paenibacillus macerans (Bacillus macerans).
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
OC NCBI_TaxID=44252;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=91109712; Pubmed=2274030;
RA Borriss R., Buettner K., Maentsaelae P.;
RT "Structure of the beta-1,3-1,4-glucanase gene of Bacillus macerans;
RL Mol. Gen. Genet. 222:278-283(1990).
RN [2]
RP ACTIVE SITE.
RX MEDLINE=93094208; Pubmed=1360982;
RA Hoel P.B., Condron R., Traeger J.C., McAnuliffe J.C., Stone B.A.;
RT "Identification of glutamic acid 105 at the active site of Bacillus
RT amyloliquefaciens 1,3-1,4-beta-D-glucan 4-glucanohydrolase using
RL epoxide-based inhibitors".
RN J. Biol. Chem. 267:25059-25066(1992).
RN [3]
RP MUTAGENESIS OF GLU-128.
RA Olsen O.;
RL Thesis (1990), University of Aarhus, Denmark.
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=93281743; Pubmed=8099449;
RA Kettel T., Simon O., Borriss R., Heinemann U.;
RT "Molecular and active-site structure of a Bacillus 1,3-1,4-beta-
RT glucanase.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:5287-5291(1993).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX MEDLINE=96028129; Pubmed=7588726;
RA Hahn M., Kettel T., Heinemann U.;
RT "Crystal and molecular structure at 0.16-nm resolution of the hybrid
RT Bacillus endo-1,3-1,4-beta-D-glucan 4-glucanohydrolase H(A16-M)".
RL Eur. J. Biochem. 233:849-858(1995).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X55959; CAA39426.1; -;
DR PIR: S11927; S11927.
DR PDB: 2AYH; 31-MAR-95.
DR PDB: 1BYH; 31-OCT-93.
DR PDB: 1CPM; 22-JUN-94.

DR PDB: 1CPN; 22-JUN-94.
 DR PDB: 1MAJ; 27-FEB-95.
 DR PDB: 1AJK; 06-MAY-98.
 DR PDB: 1AJO; 06-MAY-98.
 DR InterPro; IPR000757; Glyco_hydro_16.
 DR Pfam; PF00722; Glyco_hydro_16; 1.
 DR PRINTS; PR00737; GLHYDRLASE16.
 DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
 KW Hydrolase; Glycosidase; Signal; 3D-structure.
 FT SIGNAL 1 23
 FT CHAIN 24 237
 FT ACT_SITE 128 128
 FT ACT_SITE 132 132
 FT DISULFID 55 84
 FT MUTAGEN 128 128
 FT STRAND 26 26
 FT STRAND 29 31
 FT STRAND 41 43
 FT STRAND 51 52
 FT STRAND 53 53
 FT STRAND 56 57
 FT STRAND 59 61
 FT STRAND 62 64
 FT STRAND 66 67
 FT STRAND 70 78
 FT STRAND 79 80
 FT STRAND 81 89
 FT STRAND 93 93
 FT STRAND 96 103
 FT STRAND 108 109
 FT STRAND 110 118
 FT STRAND 120 122
 FT STRAND 123 123
 FT STRAND 127 134
 FT STRAND 135 136
 FT STRAND 137 147
 FT STRAND 148 149
 FT STRAND 150 150
 FT STRAND 155 158
 FT STRAND 163 164
 FT STRAND 168 174
 FT STRAND 179 183
 FT STRAND 184 185
 FT STRAND 186 191
 FT STRAND 200 209
 FT STRAND 213 216
 FT STRAND 225 236
 SO SEQUENCE 237 AA; 26589 MW; 436EABCDFFC87781 CRC64;
 Query Match 12.5%; Score 228; DB 1; Length 237;
 Best Local Similarity 34.3%; Pred. No. 1.5e-09;
 Matches 60; Conservative 25; Mismatches 74; Indels 16; Gaps 8;
 QY 23 TNSAKDFSGAELTYLEVOYGFKEARKMAAASGVSMFLYONGSEIADGRPWEVDI 82
 DB 75 TSSAYNFDEAEYRSTNIYGLYEVSMKPAKNTGIVSFFTYGSP--AHGTQWDEIDI 131
 QY 83 EVLGKNGSFQSNITITGKAGAKTSEKHAHVPADDAFHRTYGLGEMTPNRYRWTVGCEV 142
 DB 132 EFLGKDTTKVOFNNTYNGVGH--EKVVISLGFDAKSGFHRTYAFDMQPGYIKWYVDG-VL 187
 QY 143 RKEFGGVSNLTGTQG-LRFNLMSSES-AWVGOFDESKLPLFOFINWVKYKYYKT 195
 DB 188 KHT--ATANIPTSPGKIMNLMNGTGVDMWLSYNGAN-PLVLEYDMV---KTT 235
 RESULT 10
 GUB_PAEPO STANDARD; PRT; 238 AA.
 ID GUB_PAEPO
 AC P45797;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
 DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
 GN GLUB.
 OS Paenibacillus polymyxa (Bacillus polymyxa).
 OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
 OX NCBI_taxid=1406;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 842;
 RA MEDLINE-92041687; PubMed-1938968;
 RX Gosalbes M.J., Perez-Gonzalez J.A., Gonzalez R., Navarro A.;
 RT "Two beta-glucanase genes are clustered in Bacillus polymyxa:
 RT molecular cloning, expression, and sequence analysis of genes
 RT encoding a xylanase and an endo-beta-(1,3)-(1,4)-glucanase.";
 RL J. Bacteriol. 173:7705-7710(1991).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glucosidic linkages
 CC in beta-D-glucans containing 1,3- and 1,4-bonds.
 CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
 CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL_HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X57094; CAA40379.1; -
 DR HSSP; P23904; 1CPN.
 DR InterPro; IPR000757; Glyco_hydro_16.
 DR Pfam; PF00722; Glyco_hydro_16; 1.
 DR PRINTS; PR00737; GLHYDRLASE16.
 DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
 KW Hydrolase; Glycosidase; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 238
 FT ACT_SITE 129 129
 FT ACT_SITE 133 133
 FT DISULFID 56 85
 FT SEQUENCE 238 AA; 26919 MW; C0CF7B4E5D40E8C CRC64;
 Query Match 12.5%; Score 227.5; DB 1; Length 238;
 Best Local Similarity 34.3%; Pred. No. 1.6e-09;
 Matches 61; Conservative 26; Mismatches 74; Indels 17; Gaps 9;
 QY 20 ALTTNVAKDFSGAELTYLEVOYGFKEARKMAAASGVSMFLYONGSEIADGRPWE 79
 DB 74 SLTSPANKK-FDCEYRSTNNYGLYEVSMKPAKNTGIVSFFTYGPSH--GTQWDE 129
 QY 80 VDIEVLKNGSFQSNITITGKAGAKTSEKHAHVPADDAFHRTYGLGEMTPNRYRWTVG 139
 DB 130 IDIFELGKDTTKVOFNNTYNGVGH--EKIINLGFPASTFHTYAFDMQPGYIKWYVDG 186
 QY 140 QEVKRTGCGVSNLTGTQG-LRFNLMSSES-AWVGOFDESKLPLFOFINWVKYKYYKT 195
 DB 187 VLKHT--ATTNIPSTPGKIMNLMNGTGVDMWLSYNGAN-PLVLEYDMV---KTT 236
 RESULT 11
 EXOK_RHIME
 ID EXOK_RHIME STANDARD; PRT; 269 AA.
 AC P33693;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Endo-1,3-1,4-beta-glucanase exok precursor (EC 3.2.1.-) (Succionglycan
 DE biosynthesis protein exok).
 GN EXOK OR R81060 OR S8B20953;
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid pSymB (megaplasmid 2).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Slnorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=94042869; PubMed=8226645;
 RA Glucksmann M.A., Reuber T.L., Walker G.C.;
 RT "Family of glycosyl transferases needed for the synthesis of
 RL succinoglycan by Rhizobium meliloti."; J. Bacteriol. 175:7033-7044(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RCR2011 / S047;
 RX MEDLINE=93241147; PubMed=8479421;
 RA Becker A., Kleickmann A., Arnold W., Puehler A.;
 RT "Analysis of the Rhizobium meliloti exoH/exoK/exoL fragment: ExoK
 RT shows homology to excreted endo-beta-1,3-1,4-glucanases and ExoH
 RL resembles membrane proteins."; Mol. Gen. Genet. 238:145-154(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396508; PubMed=11481431;
 RA Phan T.M., Weidner S., Wong K., Buhmester J., Chain P.,
 RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Couzy J.,
 RA Golding B., Puehler A.;
 RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
 RT fixing endosymbiont Slnorhizobium meliloti."; Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=9826741; PubMed=9560202;
 RA York G.M., Walker G.C.;
 RT "The Rhizobium meliloti ExoK and ExoH glycanases specifically
 RT depolymerize nascent succinoglycan chains."; Proc. Natl. Acad. Sci. U.S.A. 95:4912-4917(1998).
 CC -1- FUNCTION: CLEAVES HIGH MOLECULAR WEIGHT SUCCINOGLYCAN TO YIELD LMW
 CC SUCCINOGLYCAN. DYNAMICALLY REGULATE THE MOLECULAR WEIGHT
 CC DISTRIBUTION OF SUCCINOGLYCAN BY CLEAVING NASCENT SUCCINOGLYCAN
 CC ONLY DURING A LIMITED PERIOD AFTER ITS SYNTHESIS, PERHAPS BEFORE
 CC IT UNDERGOES A TIME-DEPENDENT CHANGE IN ITS CONFORMATION OR
 CC AGGREGATION STATE.
 CC -1- PATHWAY: Exopolysaccharide biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L20758; AAA16048.1; -;
 DR EMBL: Z17219; CAA78927.1; -;
 DR EMBL: AL603645; CAC49480.1; -;
 DR HSSP: P23904; IAKK.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16.
 DR PRINTS: PR00737; GLHYDRLASE16.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 KM Exopolysaccharide synthesis; Glycosidase; Hydrolase; Signal; Plasmid;
 FT SIGNAL
 FT CHAIN 1 27 POTENTIAL.
 FT CONFLICT 28 269 ENDO-1,3-1,4-BETA-GLYCANAASE EXOK.
 FT CONFLICT 93 102 ACCEIQRKR -> LRRPDAQG (IN REF. 2).
 SQ SEQUENCE 269 AA; 30083 MW; 073CC7ED5EF2611 CRC64;

Query Match 10.8%; Score 197; DB 1; Length 269;
 Best Local Similarity 32.5%; Pred. No. 2.8e-07;
 Matches 50; Conservative 31; Mismatches 61; Indels 12; Gaps 6;

OY 25 VSADFSGAELTYLLEVOYGFKEARMKMAASGTYSWFLYONGSEIADGRPWVEVDIEV 84
 DB 87 VKERNFACGEIOTRFRFCYGYEARIKRADDSSGLNSAFYITGP---ADKRPHEIDIEV 143
 OY 85 LGRNPGSFQSN-IITGKGAOKTSEKHHAVSPADQAFHTYGLPTPYVWTVGDEVR 143
 DB 144 LGRNTAKVQINQVYSAKAGNEFLAD---VPGANGQCFNDYAFWEKRIIRIYVNGELVH 199
 OY 144 K-TEGQVSNLTGQGLRPNLMSSSSAN-WVGQF 175
 DB 200 EVTDPAKIP--VNAQKIFPSLMDTDTLDMWGTF 231
 RESULT 12
 UTR2_YEAST
 ID UTR2_YEAST STANDARD: PRT: 347 AA.
 AC P32623;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE UTR2 protein (Unknown transcript 2 protein).
 OS UTR2 OR YEL040W OR SYCP-ORF18.
 GN Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,
 RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
 RL Submitted (FEB-1993) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunkeler-Smith S.,
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 RA Petrel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Borstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B-6441;
 RX MEDLINE=94016558; PubMed=8411151;
 RA Melnick L., Sherman F.;
 RT "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,
 RT of Saccharomyces cerevisiae share a common ancestry."; J. Mol. Biol. 233:372-388(1993).
 CC -1- SIMILARITY: SOME, TO YEAST YGR189C.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U18779; AAB65002.1; ALT_INIT.
 DR EMBL: L22173; AAA34941.1; -;
 DR EMBL: S65964; AAD13975.1; -;
 DR EMBL: S66130; AAB28444.1; -;
 DR PIR: S30839; S30839.
 DR HSSP: P23904; IAKO.
 DR SCD: S0000766; UTR2.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 FT DOMAIN 234 322 SER-RICH.
 FT DOMAIN 269 283 POLY-SER.
 FT CONFLICT 10 10 L -> V (IN REF. 3).
 FT CONFLICT 171 171 A -> R (IN REF. 3).


```
CC      OPTIMUM OF 85 DEGREES CELSIUS AND A PH OPTIMUM OF 7.0 .
CC      -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC      In beta-D-glucans containing 1,3' and 1,4'bonds.
CC      -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U04836; AAA60459.1;
DR      InterPro: IPR000757; Glyco_hydro_16.
DR      Pfam; PF00722; Glyco_hydro_16; 1.
DR      PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
RW      Hydrolase; Glycosidase; Signal.
FT      SIGNAL             1          30
FT      CHAIN              31         286
FT      ACCT_SITE          158        158    NUCLEOPHILE (BY SIMILARITY).
FT      ACCT_SITE          163        163    PROTON DONOR (BY SIMILARITY).
SQ      SEQUENCE   286 AA;  33145 MW;  7215C3B624135191 CRC64;
Query Match                               7.3%; Score 133.5; DB 1; Length 286;
Best Local Similarity 23.7%; Pred. No. 0.01;
Matches 44; Conservative 34; Mismatches 85; Indels 23; Gaps 8;
QY      28 KDEFGAELTYLEEVO--YGRFEARKMKAASGTVSSNFIYONSGEINDGRWV---ENDI 82
Db       103 REYTSARLYTRGKASKMTYGRFETLRPLPSGGRTPALIMLPDRQTYSAY-WPDNGEIDI 161
           :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      83 -EVLGKNP----GSFSOINILTKAGAAQKTSEKHNHAVSPADQAHTGTGLETPNYVRVTY 137
Db       162 MEHVGNDDVVHGTVHKRAYNHLLTGROGGSIR--VPARTDRHYVAIEWTEPEIRMEY 218
           ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      QY      138 DGOEVRKTEGGGVSNLTG-----TQGLRPNLMSSSESAMVGQ-FDESKLPFLPTINW 188
Db       219 DDLSLYRRPNERLTDPEADMRHPDPDFPHLLIMIAVGAWGGQGDVPEAFAPQLVVYD 278
           |:::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      QY      189 VKVKRY 194
Db       Db       279 VRVYRW 284
           |:::|:::|
RESULT 15
BRU1_SOYBN BRU1_SOYBN STANDARD; PRT; 283 AA.
ID ID BRU1_SOYBN STANDARD; PRT; 283 AA.
AC P35694;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Brassinosteroid-regulated protein BRU1 precursor.
OS Glycine max (Soybean) .
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudictots; Rosidae;
OC eurosids II; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Epicotyl;
RC MEDLINE=94159788; PubMed=8115544;
RA Zurek D.M., Clouse S.D.;
RT "Molecular cloning and characterization of a brassinosteroid-regulated
RT plant gene elongating soybean (Glycine max L.) epicotyls.";
RL Gene Physiol. 104:161-170(1994);
CC -1- FUNCTION: POSSIBLE ROLE IN BRASSINOSTEROID-STIMULATED ELONGATION.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
```

CC modifies and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: L22162; AAA81350.1; -
DR HSSP: P23904; 1A7K
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16_1
DR Prosite: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Glycosylase; Hydrolase; Signal.
FT SIGNAL 1
FT CHAIN 31
FT CDSIN 283
SEQUENCE 263 AA: 32254 MW: 64248610 Da
BRASSINOSTEROID-REGULATED PROTEIN BRU1.
C428610E0C7835737 CCK64;

Query Match:	7.0%;	Score 128;	DB 1;	length 283;
Best Local Similarity:	23.7%;	Pred. No.	0.026;	
Matches	46;	Conservative	41;	Mismatches 73;
			Indels	34;
			Gaps	9;

```

OY 27 AKDSGALTYL-----EEVOYGFEEABMKAA--ASOTVSMFYONGSEI 71
Db 45 AKIFNGGOLLSDKDVSGSGFKSKKEYLFGIDMLKLVAAGNSAGTYAYIL-----S 98

OY 72 ADGRWWEVDIEVLK---NPGSFOSNITTKAGAOKTSEKNHNAVSPAAQOAFHTGLEM 128
Db 99 SQGRHDEIDREFELCNLSDPYLIHTNIFTOQKG--NREQOFLWEPDPTN--FHTYSIIM 155

OY 129 TPNYVKRYTVDCEVAKTBCGOVSNL--TGTOGLRF--NIMSSESAAMVG--QFDESKLP 181
Db 156 KPOHIEFLVDMPTRIEVFKNAEPLGPPFPKNQPMRIYSSLMNADDMATRGVLKTDWSKAP 215

OY 182 LEOFLMWKYVKYT 195
Db 216 FTAYYRNFKALEFS 229

```

```
Search completed: July 11, 2003, 10:53:10
Job time : 13 secs .
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:06:42 ; Search time 6.3181 Seconds
(without alignments)
1628.889 Million cell updates/sec

Title: US-09-654-652A-1

Perfect score: 1333
Sequence: 1 MVSANDFSGAEIYLTLEEVY.....TRKGFSEFNQVPRDDEPAP 248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1328	99.6	349	1 GUB_FIBSU	P17889 fibrobacter
2	266.5	20.0	259	1 GUB_BACBR	P37073 bacillus br
3	245.5	18.4	242	1 GUB_BACSU	P04957 bacillus su
4	242.5	18.2	802	1 XYND_RUMFL	O53117 ruminococcu
5	238.5	17.9	239	1 GUB_BACAM	P07980 bacillus am
6	237	17.8	334	1 GUB_CLOTM	P29716 clostridium
7	229.5	17.2	243	1 GUB_BACLI	P27051 bacillus li
8	226	17.0	237	1 GUB_PAEWA	P23304 paenibacill
9	225	16.9	238	1 GUB_PAEPO	P45797 paenibacill
10	197	14.8	269	1 EXOK_RHIME	P33693 rhizobium m
11	182.5	13.7	507	1 YG46_YEAST	P53301 saccharomyc
12	150.5	11.3	269	1 MER5_ARATH	P24806 arabidopsis
13	147	11.0	347	1 UTR2_YEAST	P32823 saccharomyc
14	133.5	10.0	286	1 GUB_RHOMR	P45798 rhodothermu
15	128	9.6	283	1 BRU1_SOYBN	P35694 glycyne max
16	117	8.8	397	1 CGKA_ALTCA	P43478 alteromonas
17	115	8.6	682	1 EL3B_BACCI	P23903 bacillus ci
18	105	7.9	465	1 EXSH_RHIME	O33880 rhizobium m
19	102.5	7.7	879	1 GUN1_CLOTM	O02934 clostridium
20	102.5	7.7	1580	1 ACC8_HUMAN	O09428 homo sapien
21	95.5	7.2	1581	1 ACC8_CRICK	O09429 rattus norv
22	95.5	7.2	1581	1 ACC8_RAT	O09429 rattus norv
23	94	7.1	465	1 EGLC_RHIME	O29392 rhizobium m
24	91.5	6.9	571	1 XYNC_PSEPL	P23031 pseudomonas
25	90	6.8	490	1 AICE_PSEAR	P18895 pseudomonas
26	88.5	6.6	736	1 VP4_ROTFC	P26193 porcine rot
27	87.5	6.6	400	1 GUN5_BACAG	O85465 bacillus ag
28	87	6.5	364	1 VM21_BORHE	P21875 borrelia he
29	87	6.5	411	1 DEOB_LACIA	O96112 lactococcus
30	86.5	6.5	1220	1 C5AC_BACTU	P56925 bacillus th
31	86.5	6.5	1385	1 C5AA_BACUD	O45760 bacillus th
32	86	6.5	354	1 DP42_RHIME	O92486 rhizobium m
33	84.5	6.3	409	1 GUN2_BACSA	P06565 bacillus sp

34	84.5	6.3	1289	1 C5AB_BACUD	O45753 bacillus th
35	83.5	6.3	422	1 FEM2_HUMAN	O94018 homo sapien
36	83.5	6.3	1531	1 FEM2_CHLTR	O84818 chlamydia t
37	82	6.2	422	1 FEM2_MOUSE	O60584 mus musculu
38	82	6.2	1103	1 CHS6_USTMA	O13395 usellago ma
39	81.5	6.1	614	1 BTUB_ECOLI	P06129 escherichia
40	81	6.1	411	1 DEOB_LACIC	O32808 lactococcus
41	81	6.1	551	1 AMT4_PSESA	P22963 pseudomonas
42	80.5	6.0	462	1 CBSA_SULAC	O54088 sulfolobus
43	80.5	6.0	708	1 PA11_DAUCA	O23865 daucus caro
44	79.5	6.0	571	1 PA14_POPKI	O40910 populus kit
45	79.5	6.0	2628	1 HAQA_PORGI	O51845 porphyromon

ALIGNMENTS

```
RESULT 1
GUB_FIBSU          STANDARD:      PRT:      349 AA.
ID P17989:
AC 01-NOV-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1999 (Rel. 33, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Mixed linkage beta-
DE glucanase) (Lichenase).
OS Fibrobacter succinogenes (Bacteroides succinogenes).
OC Bacteria: Fibrobacter/Acidobacteria group: Fibrobacter group:
OC Fibrobacter.
OX NCBI_TaxID=833;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 28-57.
RC STRAIN=Isolate S85;
RX MEDLINE=90299807; PubMed=2193918;
RA Teather R.M., Erile J.D.;
RT "DNA sequence of a Fibrobacter succinogenes mixed-linkage
RT beta-glucanase (1,3-1,4-beta-D-glucan 4-glucanohydrolase) gene.";
RL J. Bacteriol. 172:3837-3841(1990).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC -!- in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to: license@sib-sib.ch).
-----
CC EMBL: M3676; AAA2486.1; -
CC PIR: A4507; A4507.
CC HSSP: P23904; IANK.
CC InterPro: IPR000757; Glyco_hydro.16.
CC Pfam: PF00722; Glyco_hydro.16; 1.
CC PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
CC KEGG: Hydrolyase; Glycosidase; Signal; Repeat.
CC FT SIGNAL 1 27
CC FT CHAIN 28 349
CC FT ACT_SITE 79 79
CC FT ACT_SITE 83 83
CC FT DOMAIN 271 307
CC FT REPEAT 271 277
CC FT REPEAT 278 284
CC FT REPEAT 285 291
CC FT REPEAT 292 298
CC FT REPEAT 301 307
CC FT REPEAT 301 307
CC SO SEQUENCE 349 AA; 37737 MW; 16DC4F5BDFC578A CRC64;

Query Match          99.6%; Score 1328; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.2e-103;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Oy	2	VSAXPFSAGELTYLLEEVYOGKFEARMKMAAASGTVSSMFLCUNGSEIADGSPMVEVDIEV	61
Db	25	VSARDFSSABELTYLLEEVYOGKFEARMKMAASGTVSSMFLCUNGSEIADGSPMVEVDIEV	84
Oy	62	LGRNPGSFOSNIITGKAGAOKTSEKHHAASPADQAFHTYGLEMTPNYVRMTYDQGEVRK	121
Db	85	LGRNPGSFOSNIITGKAGAOKTSEKHHAASPADQAFHTYGLEMTPNYVRMTYDQGEVRK	144
Oy	122	TEGGVSNLTTGTQGLRFNLMSSESAAWVGQFDESKLPLFCFETNNVKKYKYTPCGGEGGSD	181
Db	145	TEGGVSNLTTGTQGLRFNLMSSESAAWVGQFDESKLPLFCFETNNVKKYKYTPCGGEGGSD	204
Oy	182	FTLQMTDNFDFPDSGRMKGGMTFDGNNVYDLTDKNIYSRQMLLTALTRKQSESENOVP	241
Db	205	FTLQMTDNFDFPDSGRMKGGMTFDGNNVYDLTDKNIYSRQMLLTALTRKQSESENOVP	264
Oy	242	RDDEPAP 248	
Db	265	RDDEPAP 271	

ID	GUB_BACBR	STANDARD.	PRT.	259 AA.
RESULT 2				
GUB_BACBR				
AC	P37073:			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	01-JUN-1994 (Rel. 29, Last annotation update)			
DE	Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)			
DE	(1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).			
GN	BGLB.			
OS	Bacillus brevis.			
OC	Bacteria: Firmicutes: Bacillales: Paenibacillaceae: Brevibacillus.			
OX	NCBI_TaxID=1393;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93159752; PubMed=7763386;			
RA	Louw M.E., Reid S.J., Watson T.G.;			
RT	"Characterization, cloning and sequencing of a thermostable			
RT	endo-(1,3-1,4) beta-glucanase-encoding gene from an alkalophilic			
RL	Bacillus brevis."			
RL	Appl. Microbiol. Biotechnol. 38:507-513(1993).			
CC	-1- FUNCTION: HYDROLYSES B-GLUCANS CONTAINING MIXED BETA-1,3 AND			
CC	BETA-1,4 LINKAGES. THIS ENZYME IS THERMOSTABLE: ITS OPTIMAL			
CC	TEMPERATURE IS NEAR 65-70 DEGREES CELSIUS.			
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages			
CC	in beta-D-glucans containing 1,3- and 1,4-bonds.			
CC	-1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE			
CC	SIMILAR TO LICHENASE OF GERMINATING BARLEY			
CC	-1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.			
CC	-----			
CC	This SWISS-PROT entry is copyrighted. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@lsb-sib.ch).			
CC	-----			
DR	EMBL; M84339; AAA22265.1; .			
DR	HSSP; P23904; IAKK.			
DR	InterPro: IPR000757; Glyco_hydro.16.			
DR	Pfam; PF00722; Glyco_hydro.16; 1.			
DR	PRINTS; PR00737; GLHYDRLASE16.			
DR	PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.			
KM	Hydrolase; Glycosidase; Signal.			
FT	SIGNAL	1	31	POTENTIAL.
FT	CHAIN	32	259	BETA-GLUCANASE.
FT	ACT_SITE	142	142	NUCLEOPHILE (BY SIMILARITY).
FT	ACT_SITE	146	146	PROTON DONOR (BY SIMILARITY).
QO	SEQUENCE	259 AA;	29960 MW;	A63C09F2B1FF5D13 C6C64;

Query Match	20.0%	Score 266.5	DB 1:	Length 259;
Best Local Similarity	36.0%	Pred. No. 2.8e-15;		
Matches	63:	Conservative	24:	Mismatches 75; Indels 13; Gaps 6;

Qy	3	SAKDSGAEALYLEVOYGKFEARKMAAASGTASMFLYQNGSETIADGRPAVEIDYL	62
	:		
Db	90	SAKRNKAGELKRNDRYHNGLEFVSKRPKAVEGTSSFFLY-IGEDMDQDPPDEIDIEL	148
Qy	63	GKNPSPFSNIIITGKAGAKOFTSEKHHAAVSPAADQAFFHTYGLEMTPNYARWTVDGGEVRKT	122
	:	:	:
Db	149	GKDTIRIOFNFTNGVG---NEFYLDLGFDASESFNFTAFEMRSDSLTWYNGEAVHTA	205
Qy	123	EGGVSNLTGTF-QGLRFNIMMSESA-AWYGGQDESKLPLFQGLINWYKYKKTTPGQ	175
Db	206	---TENIIPQTPQKIMMNLMPGVGDGATGVLGDDGNTPIYSYDDW---RYPLQ	253

RESULT 3			
ID	CUB_BACSU	STANDARD:	PRT: 242 AA.
AC	P04957.		
DT	13-AUG-1987 (Rel. 05, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)		
DE	(1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).		
CN	BGLS OR BGL OR LICS OR N15B.		
OC	Bacillus subtilis.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_taxid=1423;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C120;		
RX	MEDLINE=84272222; PubMed=6087283;		
RA	Murphy N., McConnell D.J., Cantwell B.A.;		
RT	"The DNA sequence of the gene and genetic control sites for the		
RL	excreted B. subtilis enzyme beta-glucanase.";		
RN	Nucleic Acids Res. 12:5355-5367(1984).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HL-25;		
RC	Tezuka H., Yunki T., Yabuchi S.;		
RT	"Construction of a beta-glucanase hyperproducing Bacillus subtilis		
RT	using the cloned beta-glucanase gene and a multi-copy plasmid.";		
RL	Agric. Biol. Chem. 53:2335-2339(1989).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=168 / BGSC1A1;		
RX	MEDLINE=97124196; PubMed=8969509;		
RA	Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,		
RA	Miwa Y., Fujita Y.;		
RT	"Sequencing of a 65 kb region of the Bacillus subtilis genome		
RT	containing the ltc and cel loci, and creation of a 177 kb contig		
RL	covering the gut-sacXy region.";		
RN	Microbiology 142:3113-3123(1996).		
RP	[4]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=168;		
RC	MEDLINE=95219081; PubMed=7704256;		
RA	Wolf M., Geczi A., Simon O., Boriss R.;		
RT	"Genes encoding xylan and beta-glucan hydrolising enzymes in Bacillus		
RT	subtilis: characterization, mapping and construction of strains		
RT	deficient in lichenase, cellulase and xylanase.";		
RL	Microbiology 141:281-290(1995).		
RN	[5]		
RP	SEQUENCE OF 1-6 FROM N.A.		
RC	STRAIN=BR151;		
RX	MEDLINE=96178961; PubMed=8606172;		
RA	Schneitz K., Stuelke J., Gertz S., Krueger S., Krieg M., Hecker M.,		
RA	Rak B.;		
RT	"Licit, a Bacillus subtilis transcriptional antiterminal protein of		
RT	the BglG family.";		
RL	J. Bacteriol 178:1971-1979(1996).		

[6]
RP SEQUENCE OF 29-63.
RA Yuuki T., Tezuka H., Yabuuchi S.:
RT "Purification and some properties of two enzymes from a beta-glucanase
hyperproducing strain, *Bacillus subtilis* HL-25.";
RL Agric. Biol. Chem. 53:2341-2346(1989).
CC -i- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -i- SUBCELLULAR LOCATION: Secreted.
CC -i- MISCELLANEOUS: BETA-GLUCANASES OF *BACILLUS* HAVE A SUBSTRATE RANGE
SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -i- MISCELLANEOUS: THE SEQUENCE OF STRAIN 168 IS SHOWN.
CC -i- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X00754; CA25328.1; -
DR EMBL: D00518; BAA0405.1; -
DR EMBL: D83026; BAA11697.1; -
DR EMBL: 246862; CAA86922.1; -
DR EMBL: 228340; CAA82195.1; -
DR EMBL: 299124; CAB15943.1; -
DR PIR: A22914; LXBS.
DR PIR: J00110; J00110.
DR HSSP: P27051; 16GB.
DR Subtilist; BG10476; bg1s.
DR Interpro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR PRINTS: PR00737; GLHYDRASE16.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Hydrolyase; Glycosidase; Signal; Complete proteome.
FT SIGNAL 1 28
FT CHAIN 29 242 BETA-GLUCANASE.
FT MOD_RES 29 29 PYRROLIDONE CARBOXYLIC ACID.
FT ACT_SITE 133 133 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 137 137 PROTON DONOR (BY SIMILARITY).
FT DISULFID 60 89 BY SIMILARITY.
FT VARIANT 24 24 A -> S (IN STRAIN HL-25).
FT VARIANT 83 83 A -> L (IN STRAIN HL-25).
FT VARIANT 204 204 P -> S (IN STRAIN C120).
SQ SEQUENCE 242 AA; 27268 MW; 45958DEA70F22B29 CRC64;

Query Match 18.4%; Score 245.5; DB 1; Length 242;
Best Local Similarity 35.8%; Pred. No. 1.4e-13;
Matches 59; Conservative 24; Mismatches 71; Indels 11; Gaps 6;

QY 7 FSGAELYTLEVOYGFKFRMKMAAASGVSSMFLYQNSGLADGRPWEVDIEVLGNKP 66
DB 87 FCGGERSRVQYTGGLYEVFRMKPAKNTGIVSSFFTYTGPT---DGPWEIDIEFEGKDT 143
QY 67 GSGFSLITGKGAQKTSKHHAVSPADQAFHTYGLTEPVPVVRMTVGGQVREKTEGQ 126
DB 144 TVQVQRYTYNGG---NHKRYIDLGFDANAHYHTAFDQPSIKMYVVGQ-LKHTATNQ 199
QY 127 VSNLTGTGGLRFLNLSSESA-AWVGQFDESKLPFLFOFIMWVYK 170
DB 200 IPTTGGK--IMNMLWNGTGVDMLGSLYNGVN-PLVAHYWVWVATK 241

RESULT 4
XIND_RUMFL STANDARD: PRT; 802 AA.
AC Q53317;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE xylanase/beta-glucanase precursor [Includes: Endo-1,4-beta-xylanase

DE (EC 3.2.1.8) (Xylanase); Endo-beta-1,3-1,4 glucanase (EC 3.2.1.73)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) [Lichenase]].
GN XIND.
OS Ruminooccus flavefaciens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;
OC Ruminooccus.
OX NCBI_TaxID=1265;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17;
RX MEDLINE=93259938; PubMed=8491715;
RA Flint H.J., Martin C., McPherson C.A., Daniel A.S., Zhang J.-X.;
RT "A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-
glucanase domains, encoded by the xynD gene of *Ruminooccus*
flavefaciens.";
RL J. Bacteriol. 175:2943-2951(1993).
CC -i- FUNCTION: CONTAINS TWO CATALYTIC DOMAINS WITH XYLANASE AND ENDO-
BETA-1,3-1,4 GLUCANASE ACTIVITIES.
CC -i- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic
linkages in xylans.
CC -i- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -i- PATHWAY: xylan degradation.
CC -i- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULOSE FAMILY
G (FAMILY 11 OF GLYCOSYL HYDROLASES).
CC -i- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO FAMILY 16 OF
GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S61204; AAB26520.1; -
DR HSSP: P23904; IAJK.
DR Interpro: IPR003305; CBM_Cenc.
DR Interpro: IPR001137; GH_11.
DR Pfam: PF00457; Glyco_hydro_11; 1.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR Pfam: PF02018; CBM_4_9; 1.
DR PRINTS: PR00911; GLHYDRASE11.
DR PROSITE: PR00737; GLHYDRASE16.
DR PROSITE: PS00776; GLYCOSYL_HYDROL_F11_1; 1.
DR PROSITE: PS00777; GLYCOSYL_HYDROL_F11_2; 1.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Xylan degradation; Hydrolase; Glycosidase; Signal;
KW Multifunctional enzyme.
FT SIGNAL 1 31
FT CHAIN 32 802 POTENTIAL.
FT DOMAIN 32 244 XYLANASE/BETA-GLUCANASE.
FT DOMAIN 245 523 A (XYLANASE).
FT DOMAIN 524 555 B.
FT DOMAIN 556 802 LINKER.
FT ACT_SITE 124 124 C (BETA-GLUCANASE).
FT ACT_SITE 226 226 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 684 684 PROTON DONOR (BY SIMILARITY).
FT DOMAIN 524 529 NUCLEOPHILE (BY SIMILARITY).
FT DOMAIN 532 543 POLY-THR.
FT DOMAIN 546 553 POLY-THR.
SQ SEQUENCE 802 AA; 89091 MW; 2860A689647284AF CRC64;

Query Match 18.2%; Score 242.5; DB 1; Length 802;
Best Local Similarity 34.9%; Pred. No. 1e-12;
Matches 58; Conservative 21; Mismatches 74; Indels 13; Gaps 6;

QY 7 FSGAELYTLEVOYGFKFRMKMAAASGVSSMFLYQNSGLADGRPWEVDIEVLGNKP 66
DB 634 YSGGFRTNPFYHYGYECSQAMKNDGVSSFFTYGPT---DDNPWEIDIEITLGNKT 690

```

OY 67 GSFOSNIITGKAGAKTSEKHNHVA5PADQAFHTYGLWTPNYVRWTVDOGEVRKTEGQ 126
DB 691 TQVQFNYNTGQCKH---EKLYDGLGFDSSSEAYHTYGFDMQPNVIAMVYDGEVRYRA----- 743
OY 127 VSNULTGTQC-LRPNLWSSSEA-AWVGQFDESKLPLOFINWVYVK 170
DB 744 TODIPRTPGKIMNAMPGLTVDDMLKAFN-GRPTLTAHVOMVYVK 788

RESULT 5
GUB_BACAM
ID GUB_BACAM STANDARD: PRT: 239 AA.
AC P07980:
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan-4-glucanohydrolase) (Lichenase).
GN BGLA.
OS Bacillus amyloliquefaciens.
OC Bacteria: Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BE 20/78;
RX MEDLINE=87192007; PubMed=3106158;
RA Hofmeister J., Kurtz A., Borris R., Knowles J.;
RT "The beta-glucanase gene from Bacillus amyloliquefaciens shows
RT extensive homology with that of Bacillus subtilis."
RL Gene 49:177-187(1986).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC DR EMBL: M15674; AAA87323.1; -.
CC DR PIR: A29091; A29091.
CC DR HSSP: P27051; IGBG.
CC DR InterPro: IPR000757; Glyco_hydro_16.
CC DR Pfam: PF00722; Glyco_hydro_16; 1.
CC DR PRINTS: PR00737; GLHYDRLASE16.
CC DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
CC KW Hydrolyase; Glycosidase; Signal.
CC FT SIGNAL 1 25
CC FT CHAIN 26 239
CC FT ACT_SITE 134 134
CC FT DISULFID 57 86
CC FT SEQUENCE 239 AA; 26928 MW; A76A64268A7AAAOB CRC64;
CC
CC Query Match 17.9%; Score 238.5; DB 1; Length 239;
CC Best Local Similarity 34.3%; Pred. No. 5,4e-13;
CC Matches 58; Conservative 26; Mismatches 74; Indels 11; Gaps 6;
OY 3 SAKDFGAEIYLTLEVOYGFKAARMKMAASGVSSMPLYONGSEIADGRPWVEVDIEVL 62
DB 80 SYNKFCGCGENRSYQVGTGLYEVRMKPANKTGVSSFFTYTGTG---GTPMEIDIEFL 136
OY 63 GANPGSFOSNIITGKAGAKTSEKHNHVA5PADQAFHTYGLWTPNYVRWTVDOGEVRK 122
DB 137 GKDITKRVQNPNTYNGAG---NHEKPADLGFDAANAHYTAFDQMPNSIKWYDQG-LKHT 192
OY 123 EGGVSNLITGTCGLRPNLWSSSEA-AWVGQFDESKLPLOFINWVYVK 170
DB 193 ATTQIPAPGK---IMNMLNMGTVDDMLGSLYNGVN-PIVAHYDMWRYRK 238

```

```

RESULT 6
GUB_CLOTM
ID GUB_CLOTM STANDARD: PRT: 334 AA.
AC P29716; P37074;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan-4-glucanohydrolase) (Lichenase) (Laminarinase).
GN LICB OR LAMI.
OS Clostridium thermocellum.
OC Bacteria: Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1515;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27405 / DSM 1237;
RX MEDLINE=92155194; PubMed=1740123;
RA Schimming S., Schwarz W.H., Staedenbauer W.L.;
RT "Structure of the Clostridium thermocellum gene licB and the encoded
RT beta-1,3-1,4-glucanase. A catalytic region homologous to Bacillus
RT lichenases joined to the reiterated domain of clostridial
RT cellulases."
RL Eur. J. Biochem. 204:13-19(1992).
RN [2]
RP PRELIMINARY SEQUENCE FROM N.A.
RC STRAIN=F7;
RX MEDLINE=92095946; PubMed=1755832;
RA Zverlov V.V., Laptev D.A., Tishkov V.I., Velikodvorska G.A.;
RT "Nucleotide sequence of the Clostridium thermocellum laminarinase
RT gene."
RL Biochem. Biophys. Res. Commun. 181:507-512(1991).
RN [3]
RP REVISIONS.
RA Zverlov V.V.;
RT Submitted (Aug-1993) to the EMBL/GenBank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RA Zverlov V.V., Velikodvorska G.A.;
RT "Cloning the Clostridium thermocellum thermostable laminarinase gene
RT in Escherichia coli; the properties of the enzyme thus produced."
RL Biotechnol. Lett. 12:811-816(1990).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- SUBUNIT: MAY FORM PART OF A MULTIMERIC COMPLEX (CELLULOSE).
CC -1- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS
CC WELL AS IN OTHER C.THERMOCELLUM CELLULOSE ENZYMES. THIS DOMAIN
CC MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC DR EMBL: X63355; CAA44959.1; -.
CC DR EMBL: X58392; CAA41281.1; -.
CC DR PIR: S23498; S23498.
CC DR PIR: J50611; J50611.
CC DR PIR: S18726; S18726.
CC DR HSSP: P23904; IAJK.
CC DR InterPro: IPR002105; Dockerin_1.
CC DR InterPro: IPR002048; EF-hand.
CC DR InterPro: IPR000757; Glyco_hydro_16.
CC DR Pfam: PF00404; Dockerin_1; 2.
CC DR Pfam: PF00722; Glyco_hydro_16; 1.
CC DR PRINTS: PR00737; GLHYDRLASE16.

```


DB 197 ATTOIPOTPK--IMNMWAGVDEWLGSYN-GVTPPLSHLWRYTK 242

RESULT 8

GUB_PAEMA STANDARD: PRT: 237 AA.

AC P23904:

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)

OS Paenibacillus macerans (Bacillus macerans).

OC Bacteria: Firmicutes; Bacilliales; Paenibacillaceae; Paenibacillus.

OX NCBI_TaxID=44252;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-91109712; PubMed-2274030;

RA Boriss R., Buettner K., Maentzel P.;

RT "Structure of the beta-1,3-1,4-glucanase gene of Bacillus macerans: homologues to other beta-glucanases."

RT Mol. Gen. Genet. 222:278-283(1990).

RN [2]

RP ACTIVE SITE.

RX MEDLINE-93094208; PubMed-1360982;

RA Hoef P.B., Condron R., Traeger J.C., McAuliffe J.C., Stone B.A.;

RT "Identification of glutamic acid 105 at the active site of Bacillus amyloliquefaciens 1,3-1,4-beta-D-glucan 4-glucanohydrolase using epoxide-based inhibitors."

RT J. Biol. Chem. 267:25059-25066(1992).

RN [3]

RP MUTAGENESIS OF GLU-128.

RA Olsen O.;

RT Thesis (1990), University of Aarhus, Denmark.

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).

RX MEDLINE-93281743; PubMed-8099449;

RA Kettel T., Simon O., Boriss R., Heinemann U.;

RT "Molecular and active-site structure of a Bacillus 1,3-1,4-beta-glucanase."

RT Proc. Natl. Acad. Sci. U.S.A. 90:5287-5291(1993).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).

RX MEDLINE-96028129; PubMed-7588726;

RA Hahn M., Kettel T., Heinemann U.;

RT "Crystal and molecular structure at 0.16-nm resolution of the hybrid Bacillus endo-1,3-1,4-beta-D-glucan 4-glucanohydrolase H(Al6-M)."; Eur. J. Biochem. 232:849-858(1995).

CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages in beta-D-glucans containing 1,3- and 1,4-bonds.

CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE SIMILAR TO LICHENASE OF GERMINATING BARLEY.

CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: X55959; CAA39426.1; -

CC PIR: S11927; S11927.

DR PDB: 2AVH; 31-MAR-95.

DR PDB: 1BTH; 31-OCT-93.

DR PDB: 1CPM; 22-JUN-94.

DR PDB: 1CJN; 22-JUN-94.

DR PDB: 1MAC; 27-FEB-95.

DR PDB: 1AJK; 06-MAY-98.

DR PDB: 1AIO; 06-MAY-98.

DR InterPro: IPR000757; Glyco_hydro_16.

DR Pfam: PF00722; Glyco_hydro_16; 1.

DR PRINTS: PR00737; GLHYDRLASE16.

DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.

KW Hydrolase; Glycosidase; Signal; 3D-structure.

FT SIGNAL 1 23

FT CHAIN 24 237

FT ACT_SITE 128 128

FT ACT_SITE 132 132

FT DISULFID 55 84

FT MUTAGEN 128 128

FT STRAND 26 26

FT STRAND 29 31

FT STRAND 41 43

FT TURN 51 52

FT STRAND 53 53

FT STRAND 56 57

FT STRAND 59 61

FT HELIX 62 64

FT STRAND 66 67

FT TURN 70 78

FT STRAND 79 80

FT TURN 81 89

FT STRAND 93 93

FT STRAND 96 103

FT TURN 108 109

FT STRAND 110 118

FT STRAND 120 122

FT HELIX 123 123

FT STRAND 127 134

FT TURN 135 136

FT STRAND 137 147

FT TURN 148 149

FT STRAND 150 150

FT STRAND 155 158

FT TURN 163 164

FT STRAND 168 174

FT STRAND 179 183

FT TURN 184 185

FT STRAND 186 191

FT STRAND 200 209

FT TURN 213 216

FT STRAND 225 236

SQ SEQUENCE 237 AA; 26589 MW; 436EABCDFFC87781 CRC64;

Query Match. 17.0%; Score 226; DB 1; Length 237;

Best local Similarity 35.1%; Pred. No. 5,9e-12;

Matches 59; Conservative 23; Mismatches 70; Indels 16; Gaps 8;

QY 7 FSGAEYLTLEVOYGFEPARKMAASGVVSMFLYONGSEIADGRPMVVDIEVLGNP 66

DB 82 FDCAEYRSTNIYGVGLYEVSMKPAKMTGIYSSFFTYTGP---AHGTQWDEIDIEFLGKDT 138

QY 67 GSFGSNIITGKAGAQGTSEKHNHVAAPADQAFHTYGLTLPNVYRVRTVPGQEVKTEGGO 126

DB 139 TKVQENYVYNGVGH---EKVISLGFDAKSGFHTVAFDMPQGIKKYVVG--VLKHT---A 191

QY 127 VSNLTGTGOG-LRFNLWSSESA-AWVGQFDESKLPDFQFINWVYVRYT 172

DB 192 TANIPSTPKIMNMWAGVDEWLGSYNGAN-PLVAYEDWV---KYT 235

RESULT 9

GUB_PAEPD STANDARD: PRT: 238 AA.

AC P45797:

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)

OS Paenibacillus polymyxa (Bacillus polymyxa).

OC Bacteria: Firmicutes; Bacilliales; Paenibacillaceae; Paenibacillus.

OX NCBI_TaxID=1406;

```

RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 842;
RX MEDLINE=92041687; PubMed=1938968;
RA Gosalbes M.J., Perez-Gonzalez J.A., Gonzalez R., Navarro A.;
RT "Two beta-glycanase genes are clustered in Bacillus polymyxa:
RT molecular cloning, expression, and sequence analysis of genes
RT encoding a xylanase and an endo-beta-(1,3)-(1,4)-glucanase.";
RL J. Bacteriol. 173:7705-7710(1991).
CC -i- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -i- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC SIMILAR TO LICHEMATING BARLEY.
CC -i- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
CC EMBL: X57094; CAA40379.1; -.
CC HSSP: P23904; ICPN.
CC InterPro: IPR000757; Glyco_hydro.16.
CC Pfam: PF00722; Glyco_hydro.16; 1.
CC PRINTS: PR00737; GLHYDRLASE16.
CC PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
CC Hydrolase; Glycosidase; Signal.
CC SIGNAL 1 26
CC CHAIN 27 238
CC FT ACT_SITE 129 129 BETA-GLUCANASE.
CC FT ACT_SITE 133 133 NICLEOPHILE (BY SIMILARITY).
CC FT DISULFID 56 85 PROTON DONOR (BY SIMILARITY).
CC FT DISULFID 85 85 BY SIMILARITY.
CC SQ SEQUENCE 238 AA; 26919 MW; C0CF7B4EASD40E8C CRC64;

Query Match
Best Local Similarity 34.5%; Score 225; DB 1; Length 238;
Matches 58; Conservative 23; Mismatches 71; Indels 16; Gaps 8;

OY 7 FSGAELTYLEVOYQGFARMKMAASGVSMFLYQNGSEIADGRPWVEDIEVLGKNP 66
DB 83 FCGEGRSTNNYGVGLYEVMKPAKNTGIYSSFFTYTGPSH---GTQWDEIDIEFLGKDT 139
OY 67 GSFQSIITGKGAQCTSKKHAVSPAADAFTTYGLETPNTVVRVTVGQEVKRTESGO 126
DB 140 TVYQFVYTYNGVGH---EKIINLGFDASTSFHTYAFDMQPGYIKWYVDG-VLKHT---A 192

OY 127 VSNLGTGQGLRFNLMSSESA-AWVGQFDESKLPLEQFINMKVYKYTT 172
DB 193 TTNIPSTPGKIMNLMNGCVSWLGSYNGAN-PLTAEYDWM---KYT 236

RESULT 10
EXOK_RHIME
ID EXOK_RHIME STANDARD: PRT; 269 AA.
AC P33693;
DT 01-FEB-1994 (rel. 28, Created)
DT 01-FEB-1994 (rel. 28, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Endo-1,3-1,4-beta-glycanase exok precursor (EC 3.2.1.-) (Succinoglycan
DE biosynthesis protein exok).
GN EXOK OR RB1080 OR SMB20955.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymB (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
ON NCBI_TaxId=382;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=94042869; PubMed=8226645;

```

```

RA Glucksmann M.A., Reuber T.L., Walker G.C.;
RT "Family of glycosyl transferases needed for the synthesis of
RT succinoglycan by Rhizobium meliloti.";
RL J. Bacteriol. 175:7033-7044(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RX MEDLINE=93241147; PubMed=8479421;
RA Becker A., Kleckmann A., Arnold W., Puhler A.;
RT "Analysis of the Rhizobium meliloti exoH/exoK/exoL fragment: ExoK
RT shows homology to excreted endo-beta-1,3-1,4-glucanases and ExoH
RT resembles membrane proteins.";
RL Mol. Gen. Genet. 238:145-154(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorholter F.J., Hernandez-Lucas I., Becker A., Gouzy J.,
RA Golding B., Puhler A.;
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=98226741; PubMed=9560202;
RA York G.M., Walker G.C.;
RT "The Rhizobium meliloti ExoK and ExoH glycanases specifically
RT depolymerize nascent succinoglycan chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4912-4917(1998).
CC -i- FUNCTION: CLEAVES HIGH MOLECULAR WEIGHT SUCCINOGLYCAN TO YIELD LMW
CC SUCCINOGLYCAN. DYNAMICALLY REGULATE THE MOLECULAR WEIGHT
CC DISTRIBUTION OF SUCCINOGLYCAN BY CLEAVING NASCENT SUCCINOGLYCAN
CC ONLY DURING A LIMITED PERIOD AFTER ITS SYNTHESIS, PERHAPS BEFORE
CC IT UNDERGOES A TIME-DEPENDENT CHANGE IN ITS CONFORMATION OR
CC AGGREGATION STATE.
CC -i- PATHWAY: Exopolysaccharide biosynthesis.
CC -i- SUBCELLULAR LOCATION: Secreted.
CC -i- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
CC EMBL: L20758; AAA16048.1; -.
CC DR EMBL: Z17219; CAA78927.1; -.
CC DR EMBL: AL603645; CAC49480.1; -.
CC HSSP: P23904; IAKR.
CC InterPro: IPR000757; Glyco_hydro.16.
CC Pfam: PF00722; Glyco_hydro.16; 1.
CC PRINTS: PR00737; GLHYDRLASE16.
CC PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
CC Exopolysaccharide synthesis; Glycosidase; Hydrolase; Signal; Plasmid;
CC Complete proteome.
CC SIGNAL 1 27
CC CHAIN 28 269
CC FT CONFLICT 93 102 ACCELOTERR -> LRRNPDAG (IN REF. 2).
CC SQ SEQUENCE 269 AA; 30083 MW; 073CCTED65EFP611 CRC64;

Query Match
Best Local Similarity 14.8%; Score 197; DB 1; Length 269;
Matches 50; Conservative 31; Mismatches 61; Indels 12; Gaps 6;

OY 2 VSKQDSGAELTYLEVOYQGFARMKMAASGVSMFLYQNGSEIADGRPWVEDIEV 61
DB 87 VKERNFACGEIOTRRREGYGYEARIKADGSGLNAPFTYIGP---ADKKRHDEIDFEV 143
OY 62 LCKNPGSFQSN-IINGKAAQKTSSEKHNAVSPAADAFTTYGLETPNTVVRVTVGQEV 120

```


Db 144 LCKNTAKVOINDOYVSAKGNELAD----VPGANGFNDAFVMEKRIYYVNGELVH 199
 Oy 121 K-TEGGVSNLTGTGLRPNLWSSESA--WVGO 152
 Db 200 EYTDPAKIP--VNAOKIFSLMGTDTLTWMTGF .231

RESULT 11

YC46_YEAST STANDARD: PRT: 507 AA.

AC P53301.

DT 01-OCT-1996 (rel. 34, Created)

DT 01-OCT-1996 (rel. 34, Last sequence update)

DT 01-NOV-1997 (rel. 35, Last annotation update)

DE Hypothetical 52.8 kDa protein in BUHL-HIP1 Intergenic region.

GN YC46_YEAST OR G7553.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

NCBI_TaxID=4932;

CC [1]

CC SEQUENCE FROM N.A.

RC STRAIN=5288C;

RA MEDLINE=91279231; PubMed=9133739;

RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,

RA Nombela C.;

RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm

RL Yeast 13:357363(1997).

CC -1- SIMILARITY: SOME, TO YEAST UTR2.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: 272974; CAA67215.1;

DR EMBL: X99074; CAA67525.1;

DR HSSP: P23904; IATK.

DR SGD: S0003421; CRH1.

RESULT 12

MERS_ARATH STANDARD: PRT: 269 AA.

AC P24806; 039148; 041904; 064955.

DT 01-MAR-1992 (rel. 21, Created)

DT 15-JUN-2002 (rel. 41, Last sequence update)

DT 15-JUN-2002 (rel. 41, Last annotation update)

DE MERS-5 protein precursor (Endo-xyloglucan transferase) (Xyloglucan

DE endo-1,4-beta-D-glucanase).

GN MERS-5 OR MERS5B OR SEM OR ATAG30270 OR F9M11.120.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;

OC Eurosids II; Brassicales; Brassicaceae; Arabidopsids.

NCBI_TaxID=3702;

CC [1]

CC SEQUENCE FROM N.A.

RC MEDLINE=93005704; PubMed=1840916;

RA Medford J.I., Elmer J.S., Klee H.J.;

RT "Molecular cloning and characterization of genes expressed in shoot

RT apical meristems."

RL Plant Cell 3:359-370(1991).

CC [2]

CC SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RA Kaminal T., Tomita E., Nishitani K.;

RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.

CC [3]

CC SEQUENCE FROM N.A.

RC MEDLINE=20083488; PubMed=10617198;

RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,

RA Pohl T., Duesterhoeft A., Stiekema W., Entlan K.-D., Terryn N.,

RA Harris B., Ansoorge W., Brandt P., Griwell L.A., Rieger M.,

RA Weichselgartner M., de Simone V., Obermayer B., Meche R., Mueller M.,

RA Kreis M., Delseny M., Puldomenech P., Watson M., Schmidheini T.,

RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,

RA Vos P., Hehseisel J., Zimmermann W., Wedler H., Ridley P.,

RA Langham S.-A., McCullagh B., Billam L., Robben J.,

RESULT 13

MERS_ARATH STANDARD: PRT: 269 AA.

AC P24806; 039148; 041904; 064955.

DT 01-MAR-1992 (rel. 21, Created)

DT 15-JUN-2002 (rel. 41, Last sequence update)

DT 15-JUN-2002 (rel. 41, Last annotation update)

DE MERS-5 protein precursor (Endo-xyloglucan transferase) (Xyloglucan

DE endo-1,4-beta-D-glucanase).

GN MERS-5 OR MERS5B OR SEM OR ATAG30270 OR F9M11.120.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;

OC Eurosids II; Brassicales; Brassicaceae; Arabidopsids.

NCBI_TaxID=3702;

CC [1]

CC SEQUENCE FROM N.A.

RC MEDLINE=93005704; PubMed=1840916;

RA Medford J.I., Elmer J.S., Klee H.J.;

RT "Molecular cloning and characterization of genes expressed in shoot

RT apical meristems."

RL Plant Cell 3:359-370(1991).

CC [2]

CC SEQUENCE FROM N.A.

RC MEDLINE=20083488; PubMed=10617198;

RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,

RA Pohl T., Duesterhoeft A., Stiekema W., Entlan K.-D., Terryn N.,

RA Harris B., Ansoorge W., Brandt P., Griwell L.A., Rieger M.,

RA Weichselgartner M., de Simone V., Obermayer B., Meche R., Mueller M.,

RA Kreis M., Delseny M., Puldomenech P., Watson M., Schmidheini T.,

RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,

RA Vos P., Hehseisel J., Zimmermann W., Wedler H., Ridley P.,

RA Langham S.-A., McCullagh B., Billam L., Robben J.,

RA Van der Schueren J., Grymonprez B., Chung Y.-J., Vandenbussche F.,

RA Braeken M., Welkens I., Voet M., Bastiaens I., Aert R., Defoor E.,

RA Welzenegger T., Bothe G., Ransperger U., Hilbert H., Braun M.,

RA Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W.,

RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koelter P.,

RESULT 14

MERS_ARATH STANDARD: PRT: 269 AA.

AC P24806; 039148; 041904; 064955.

DT 01-MAR-1992 (rel. 21, Created)

DT 15-JUN-2002 (rel. 41, Last sequence update)

DT 15-JUN-2002 (rel. 41, Last annotation update)

DE MERS-5 protein precursor (Endo-xyloglucan transferase) (Xyloglucan

DE endo-1,4-beta-D-glucanase).

GN MERS-5 OR MERS5B OR SEM OR ATAG30270 OR F9M11.120.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;

OC Eurosids II; Brassicales; Brassicaceae; Arabidopsids.

NCBI_TaxID=3702;

CC [1]

CC SEQUENCE FROM N.A.

RC MEDLINE=93005704; PubMed=1840916;

RA Medford J.I., Elmer J.S., Klee H.J.;

RT "Molecular cloning and characterization of genes expressed in shoot

RT apical meristems."

RL Plant Cell 3:359-370(1991).

CC [2]

CC SEQUENCE FROM N.A.

RC MEDLINE=20083488; PubMed=10617198;

RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,

RA Pohl T., Duesterhoeft A., Stiekema W., Entlan K.-D., Terryn N.,

RA Harris B., Ansoorge W., Brandt P., Griwell L.A., Rieger M.,

RA Weichselgartner M., de Simone V., Obermayer B., Meche R., Mueller M.,

RA Kreis M., Delseny M., Puldomenech P., Watson M., Schmidheini T.,

RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,

RA Vos P., Hehseisel J., Zimmermann W., Wedler H., Ridley P.,

RA Langham S.-A., McCullagh B., Billam L., Robben J.,

RA Van der Schueren J., Grymonprez B., Chung Y.-J., Vandenbussche F.,

RA Braeken M., Welkens I., Voet M., Bastiaens I., Aert R., Defoor E.,

RA Welzenegger T., Bothe G., Ransperger U., Hilbert H., Braun M.,

RA Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W.,

RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koelter P.,

RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C.,
 RA Antonou B., Zidanic M., Strong C., Sun H., Ianar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
 RA Granaat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana";
 RL Nature 402:769-777 (1999).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "RITEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SSP consortium (Salk/Stanford/PGEC)";
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 1-132 FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=Green siliques;
 RA Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delzeny M.;
 RL Submitted (OCT-1992) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 1-120 FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=Leaf;
 RA MEDLINE=98278374; PubMed=9617812;
 RX Park J.-H., Oh S.A., Kim Y.H., Moo H.R., Nam H.G.;
 RT "Differential expression of senescence-associated mRNAs during leaf
 RT senescence induced by different senescence-inducing factors in
 RT Arabidopsis";
 RL Plant Mol. Biol. 37:445-454 (1998).
 CC [8]
 CC -1- FUNCTION: Involved in cell wall reconstruction.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN SHOOT APICAL MERISTEMS, ALSO
 CC FOUND IN SEEDLINGS AND MERISTEMS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
 CC in positions 158; 178; 183; 189; 190; 194 and 199.
 CC -1- CAUTION: Ref.6 sequence differs from that shown due to frameshifts
 CC in positions 93 and 104.
 CC [9]
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC [10]
 CC EMBL: M63166; AAA32828.1; ALT_FRAME.
 DR EMBL: D63508; BAA09783.1; -;
 DR EMBL: AL109796; CAB52471.1; -;
 DR EMBL: AL161576; CAB81020.1; -;
 DR EMBL: AY05156; AAK59660.1; -;
 DR EMBL: AY063027; AAL34201.1; -;
 DR EMBL: Z17602; CA79012.1; ALT_FRAME.
 DR EMBL: AF035384; AAC39467.1; -;
 DR PIR: J01022; J01022.
 DR HSSP: P23904; IATK.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 KW Glycosidase; Hydrolase; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 269 MERI-5 PROTEIN.
 FT CONFLICT 93 94 GS-> DR (IN REF. 7).
 FT CONFLICT 184 184 A-> G (IN REF. 1).
 SO SEQUENCE 269 AA; 30755 MW; 648F042BC7ADE86 CRC64;

Query Match 11.3%; Score 150.5; DB 1; Length 269;
 Best Local Similarity 26.3%; Pred. No. 1.4e-05;
 Matches 47; Conservative 34; Mismatches 65; Indels 33; Gaps 8;

Db 54 SCSGFSKTEYLEFKIDWQIKLVPNGSACTVTTFYLSKSGS-----TWDEIDPEFLGNM 107
 Oy 65 --NPGSFQNIITTKRGAQAKTSEKHAAVSPADQAFHTYGLMTPTMYWYVDGQEVKRT 122
 Db 108 SDDPYLHNVYTOGKG-DKEOQFLHMPPTAN--FHTYSILMNQRIILTYDDPPIREF 164
 Oy 123 EGGVSNLTCGTOGLFF-----NLWSSSAWVG---QEDSKLPLFQFIWVKV 168
 Db 165 KWE-----SLGVLPKPKPMRMATSLWNADWATRGGLVKTDSKAFPMASYNRIKI 217
 RESULT 13
 UTR2_YEAST
 ID UTR2_YEAST STANDARD; PRT; 347 AA.
 AC P32623.
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE UTR2 protein (unknown transcript 2 protein).
 GN UTR2 OR YEL040W OR SYGP-ORE18.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,
 RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
 RL Submitted (FEB-1993) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
 RA Hymen R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yellon M., Bostein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B-6441.
 RX MEDLINE=94016558; PubMed=8411151;
 RT Melnick L., Sherman F.;
 RT "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,
 RT of Saccharomyces cerevisiae share a common ancestry.";
 RL J. Mol. Biol. 233:372-388 (1993).
 CC [4]
 CC -1- SIMILARITY: SOMF, NO YEAST YGRI189C.
 CC [5]
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC [6]
 CC EMBL: U18779; AAB65002.1; ALT_INIT.
 DR EMBL: I22173; AAA34941.1; -;
 DR EMBL: S65964; AAD13975.1; -;
 DR EMBL: S66130; AAB23444.1; -;
 DR PIR: S30839; S30839.
 DR HSSP: P23904; IAOO.
 DR SGD: S0000766; UTR2.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR DOMAIN 234 322 SER-RICH.
 FT DOMAIN 269 283 POLY-SER.
 FT CONFLICT 10 10 L-> V (IN REF. 3).
 FT CONFLICT 171 171 A-> R (IN REF. 3).
 FT CONFLICT 234 234 S-> C (IN REF. 3).

Db 99 S Q S P T H D E I D E F E L G N L S G D P Y I L T H T I F I Q O G K G - N E E Q O F I Y L M E D P T R N - P H Y S I I T M 155

Qy 106 T P N Y K R W T V D Q G E V R K E T G E G O V S N L - - T G O G L R F - N L M S S S A A M V G - - O F D S K L P 158

Db 156 K P O H I E L Y D M T P I R V R N N A P L G V A P P P K O P M R I Y S I L M N A D M A T R G G L V A K T W D S K A P 215

Qy 159 L E O F I N M V K Y K Y K T 172

Db 216 F T A Y V R N F K A L E F S 229

Search completed: January 9, 2003, 12:16:46
Job time : 7.31481 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:06:42 ; Search time 6.79861 Seconds
(without alignments)
1628.889 Million cell updates/sec

Title: US-09-654-652a-2

Perfect score: 1439

Sequence: 1 MVSADFGSCALYLETLEYVY.....PNSSVDKLAALNNNNNN 267

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1340	93.1	349	1 GUB_FIBSU	P17989 fibrobacter
2	266.5	18.5	259	1 GUB_BACBR	P37073 bacillus br
3	245.5	17.1	242	1 GUB_BACSU	P04957 bacillus su
4	242.5	16.9	802	1 XYND_RUMFL	053317 rummocoocu
5	238.5	16.6	239	1 GUB_BACAM	P07980 bacillus am
6	237	16.5	334	1 GUB_CLODM	P29716 clostridium
7	229.5	15.9	243	1 GUB_BACLI	P27051 bacillus li
8	226	15.7	237	1 GUB_PAEPA	P23904 paenibacill
9	225	15.6	238	1 GUB_PAEPO	P45797 paenibacill
10	197	13.7	269	1 EXOK_RHIME	P33693 rhizobium m
11	182.5	12.7	507	1 YG46_YEAST	P53301 saccharomyc
12	150.5	10.5	269	1 MERS_ARATH	P24806 arabidopsis
13	147	10.2	347	1 UTR2_YEAST	P32623 saccharomyc
14	133.5	9.3	286	1 GUB_RHOMR	P45798 rhodothermu
15	128	8.9	283	1 BRU1_SOYBN	P35694 glycine max
16	117	8.1	397	1 CGKA_ALTCA	P43478 alteromonas
17	115	8.0	682	1 E13B_BACCI	P23603 bacillus ci
18	105	7.3	465	1 EXSH_RHIME	P03680 rhizobium ci
19	102.5	7.1	879	1 GUNT_CLODM	002934 clostridium
20	102.5	7.1	1580	1 ACCB_HUMAN	009428 homo sapien
21	95.5	6.6	1581	1 ACCB_CRICR	009427 cricetus cr
22	95.5	6.6	1581	1 ACCB_RAT	009429 rattus norv
23	94	6.5	465	1 EGIC_RHIME	092392 rhizobium m
24	91.5	6.4	571	1 XYNC_PSEFL	P23031 pseudomonas
25	90	6.3	490	1 ALGE_PSEAE	P18895 pseudomonas
26	88.5	6.2	736	1 VPA_ROMPC	P26193 porcine rot
27	87.5	6.1	400	1 GUMS_BACAG	085465 bacillus ag
28	87.5	6.1	1220	1 C5AC_BACTU	P56955 bacillus th
29	87.5	6.1	1385	1 C5AA_BACUD	045760 bacillus th
30	87	6.0	364	1 VM21_BORHE	P21875 borrelia he
31	87	6.0	411	1 DEOB_LACLA	096112 lactococcus
32	86.5	6.0	406	1 REN1_HUMAN	P00797 homo sapien
33	86	6.0	354	1 DPA2_RHIME	092488 rhizobium m

34	85.5	5.9	287	1 FRHB_METJA	Q60341 methanococc
35	85.5	5.9	289	1 C5AB_BACUD	Q45753 bacillus th
36	84.5	5.9	409	1 G5AB_BAC54	P06565 bacillus sp
37	84	5.8	361	1 WDS_DROME	Q9V318 drosophila
38	83.5	5.8	422	1 FBW2_HUMAN	Q90438 homo sapien
39	83.5	5.8	1531	1 PMPD_CHLTR	Q84818 chlamydia t
40	83.5	5.8	1545	1 ACC9_RAT	Q63563 rattus norv
41	83.5	5.8	1546	1 ACC9_MOUSE	P70170 mus musculu
42	83.5	5.8	1549	1 ACC9_HUMAN	Q60706 homo sapien
43	83.5	5.8	1549	1 ACC9_RABIT	P82451 oryctolagus
44	82	5.7	422	1 FBW2_MOUSE	Q60584 mus musculu
45	82	5.7	1103	1 CHS6_USTMA	O13395 ustilago ma

ALIGNMENTS

RESULT 1	ID	GUB_FIBSU	STANDARD:	PRT:	349 AA.
AC	P17989				
DT	01-NOV-1990 (Rel. 16, Created)				
DT	01-NOV-1990 (Rel. 16, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)				
DE	(1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Mixed linkage beta-glucanase) (Lichenase).				
OS	Fibrobacter succinogenes (Bacteroides succinogenes).				
OC	Bacteria; Fibrobacter/Acidobacteria group; Fibrobacter group;				
OC	Fibrobacter.				
OX	NCBI_TaxID=833;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 28-57.				
RC	STRAIN=Isolate S85;				
RX	MEDLINE=90299807; PubMed=2193918;				
RA	Teather R.M., Erfle J.D.;				
RT	"DNA sequence of a Fibrobacter succinogenes mixed-linkage beta-glucanase (1,3-1,4-beta-D-glucan 4-glucanohydrolase) gene";				
RL	J. Bacteriol. 172:3637-3841(1990).				
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages in beta-D-glucans containing 1,3- and 1,4-bonds.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	or send an email to license@isb-sib.ch).				
DR	EMBL: M33676; AAA24896.1; ..				
DR	PIR: A44507; A44507.				
DR	HSSP: P23904; IAJK.				
DR	InterPro: IPR000757; Glyco_hydro_16.				
DR	Pfam: PF00722; Glyco_hydro_16; 1.				
DR	PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.				
KW	Hydrolase; Glycosidase; Signal; Repeat.				
FT	SIGNAL	1	27		
FT	CHAIN	28	349		
FT	ACT_SITE	79	79		
FT	ACT_SITE	83	83		
FT	DOMAIN	271	307		
FT	REPEAT	271	277		
FT	REPEAT	278	284		
FT	REPEAT	285	291		
FT	REPEAT	292	298		
FT	REPEAT	301	307		
SO	SEQUENCE	349 AA;	37737 MW;	16DC4F5BDEFC578A CRC64;	
Query Match	93.1%;	Score 1340;	DB 1;	Length 349;	
Best Local Similarity	99.6%;	Pred. No. 2,5e-104;			
Matches 250;	Conservative	0;	Mismatches	1;	Indels
				0;	Gaps
					0;


```

RN      [6]
RP      SEQUENCE OF 29-63.
RA      Yuki T., Tezuka H., Yabuuchi S.:
RT      "Purification and some properties of two enzymes from a beta-glucanase
RT      hyperproducing strain, Bacillus subtilis HL-25."
RL      Cell. Biol. Chem. 53:2341-2346(1989).
CC      -I- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC      in beta-D-glucans containing 1,3- and 1,4-bonds.
CC      -I- SUBCELLULAR LOCATION: Secreted.
CC      -I- SUBCELLULOSES: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC      SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC      -I- MISCELLANEOUS: THE SEQUENCE OF STRAIN 168 IS SHOWN.
CC      -I- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X00754; CAZ5328.1; -
DR      EMBL: D00518; BAA00405.1; -
DR      EMBL: D83026; BAA11697.1; -
DR      EMBL: Z46862; CAAB6922.1; -
DR      EMBL: Z28340; CAAB2195.1; -
DR      EMBL: Z99124; CAB15943.1; -
DR      PIR: A22914; LXBS.
DR      PIR: J00110; J00110.
DR      HSP: P27051; 16RBG.
DR      SUDL1: B610476; bg15.
DR      InterPro: IPR000757; Glyco_hydro_16.
DR      Pfam: PF00722; Glyco_hydro_16.1.
DR      PRINTS: PR00737; GLHYDRLASE16.
DR      PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW      Hydrolyase, glycosidase, Signal, Complete proteome.
FT      SIGNAL 1 28
FT      CHAIN 29 242
FT      MOD_RES 29 29
FT      ACT_SITE 133 133
FT      ACT_SITE 137 137
FT      DISULFID 60 89
FT      VARIANT 24 24
FT      VARIANT 83 83
FT      VARIANT 204 204
FT      VARIANT 242 27268
SO      SEQUENCE 242 AA; 27268 MW; 4595BDEA70F2B29 CRC64;

Query Match 17.1%; Score 245.5; DB 1; Length 242;
Best Local Similarity 35.8%; Pred. No. 1,86-13;
Matches 59; Conservative 24; Mismatches 71; Indels 11; Gaps 6;

OY      7 FSGAELVTLSEVQYKGFPEARMKMAASGTVSMFLYONGSELAIDGRPAVEVDIEVLGNP 66
Db      | | : : : | | : | | | | | | | | | | | | | | | | | | | | | |
OY      87 FDCGEGNSVQYLYGYLEVLRKMPKRAKNTGIVSSFFTYTPT--DQTPWDEIDIEFLGKDT 143
Db      | | : : : | | : | | | | | | | | | | | | | | | | | | | | | |
OY      67 GSPFSNITTKAGAKQKTSSEKHAASPADQAFHTYGLTFWTPNYVWYVDGQEVKRTGEGQ 126
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      144 TKYQFNYVTNAG--NNEKTLVDLGFDAANAYHTYAFPMQGNSTKIVYDQGLKHTATNQ 199
Db      | | : : : | | : | | | | | | | | | | | | | | | | | | | | | |
OY      127 VSNLTGTGGLRFLNLSSESA-AWVGQFDESKLPLEFQITNWKVYK 170
Db      | | : : : | | : | | | | | | | | | | | | | | | | | | | | | |
OY      200 IPTTPGK--IMMNLNMGTVDEWLGSYNGVN-PLVANYDWDVRYTK 241
Db      | | : : : | | : | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
XIND_RUMFL STANDARD: PRT; 802 AA.
AC 053317;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE xylanase/beta-glucanase precursor [includes: Endo-1,4-beta-xylanase

```

```

DE (EC 3.2.1.8) (Xylanase); Endo-beta-1,3-1,4 glucanase (EC 3.2.1.73)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase)}.
GN XYND.
OS Ruminococcus flavefaciens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;
CC Ruminococcus.
OX NCBI_TaxID=1265;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17;
RX EMLNt_H.J., Martin J., McPherson C.A., Daniel A.S., Zhang J.-X.;
RA "A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-
RT glucanase domains, encoded by the xynd gene of Ruminococcus
RT flavefaciens."
RL J. Bacteriol. 175:2943-2951(1993).
CC -1- FUNCTION: CONTAINS TWO CATALYTIC DOMAINS WITH XYLANASE AND ENDO-
CC BETA-1,3-1,4 GLUCANASE ACTIVITIES.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic
CC linkages in xylans.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- PATHWAY: Xylan degradation.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULOSE FAMILY
CC G (FAMILY 11 OF GLYCOSYL HYDROLASES).
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO FAMILY 16 OF
CC GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S61204; AAB26620.1; -.
DR HSSP: P23904; IAKK.
DR InterPro: IPR003305; CBM_Cenc.
DR InterPro: IPR001137; GH_11.
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00457; Glyco_hydro_11.1.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR Pfam: PF02018; CBM_4_9; 1.
DR PRINTS: PR00911; GLHYDRLASE1.
DR PRINTS: PR00737; GLHYDRLASE16.
DR PROSITE: PS00777; GLYCOSYL_HYDROL_F11_1; 1.
DR PROSITE: PS00777; GLYCOSYL_HYDROL_F11_2; 1.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Xylan degradation; Hydrolase; Glycosidase; Signal;
KW Multifunctional enzyme.
FT SIGNAL 1 31
FT CHAIN 32 802
FT DOMAIN 232 244 POTENTIAL.
FT DOMAIN 445 523 A (XYLANASE).
FT DOMAIN 524 555 B.
FT DOMAIN 556 802 LINER.
FT ACT_SITE 124 124 C (BETA-GLUCANASE).
FT ACT_SITE 226 226 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 684 684 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 684 684 NUCLEOPHILE (BY SIMILARITY).
FT DOMAIN 524 529 POLY-THR.
FT DOMAIN 532 543 POLY-THR.
FT DOMAIN 546 553 POLY-THR.
SQ SEQUENCE 802 AA; 89091 MW; 2880A689647284AF CRC64;

Query Match 16.9%; Score 242.5; DB 1; Length 802;
Best Local Similarity 34.9%; Pred. No. 1,36-12;
Matches 58; Conservative % 21; Mismatches 74; Indels 13; Gaps 6;

7 FSGSELYTLEVOVGKGFARMKMAASTGVSMFLYNGSEIACGRPEVDEYILGNP 66
Db 634 YSGGEFTNNFTYHRTGYIECSQAMKNDGVSSFFTYGPS--DDNPDEIDIDILGNT 690

```

```
OY 67 GSFQSNITGKAGCAQRTSEKHNHVASPADQAFHTYGLWTPNRYRWTVDCQEVKTEGQ 126
DB 691 TQVQFVNYTNGQSKH---EKLYDLGFDSSSEAYHTYGFDMQPNIAWYDQREVYRA----- 743
OY 127 VSNLJTGTCG-LRFNWSSESA-AWGOFPDESKLP.LFOFINWVYVK 170
DB 744 TODIPKPGKIMNAMPGLIVDDMLKAFN-GRPTPLTAHYQWTVYTK 788

RESULT 5
GUB_BACAM STANDARD: PRT: 239 AA.
ID GUB_BACAM
AC P07980;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
GN BGLA.
OS Bacillus amyloliquefaciens.
OC Bacteria: Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BE 20/78;
RX MEDLINE=87192007; PubMed=3106158;
RA Hofemeister J., Kurtz A., Borliss R., Knowles J.;
RT "The beta-glucanase gene from Bacillus amyloliquefaciens shows
  extensive homology with that of Bacillus subtilis."
  Gene 49:177-187(1986).
RL
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
  in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
  SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
CC EMBL: M15674; AAA87323.1; -.
CC DR PIR: A29091; A29091.
CC DR HSSP: P27051; 1G8G.
CC DR InterPro: IPR000757; Glyco_hydro_16.
CC DR Pfam: PF00722; Glyco_hydro_16; 1.
CC DR PRINTS: PR00737; GLHYDRASE16.
CC DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
CC KM Hydrolyase; Glycosidase; Signal.
CC FT SIGNAL 1 25
CC FT CHAIN 26 239
CC FT ACT_SITE 134 134
CC FT DISULFID 57 86
CC FT SEQUENCE 239 AA; 26928 MW; A76A64268A7AA0B CRC64;
  BETA-GLUCANASE.
  NUCLEOPHILE (BY SIMILARITY).
  BY SIMILARITY.
  BEST LOCAL SIMILARITY 34.3%; Pred. No. 6,6e-13;
  Matches 58; Conservative 26; Mismatches 74; Indels 11; Gaps 6;

Query Match 16.6%; Score 238.5; DB 1; Length 239;
Best Local Similarity 34.3%; Pred. No. 6,6e-13;
Matches 58; Conservative 26; Mismatches 74; Indels 11; Gaps 6;

OY 3 SAKDFSGAELTYLLEVOYGFKFAKRMKAASGVSSMFLYQNSSETADGRPNVEVDIEVL 62
DB 80 SYNKKFCCGNSRSQYTGYLEYRKMPAKNTGIVSSFFTYTGTE---GTPNDEIDIEFL 136
OY 63 GRNPGSFQSNITGKAGCAQRTSEKHNHVASPADQAFHTYGLWTPNRYRWTVDCQEVKRT 122
DB 137 GKDTTFVQCFNYTNGAG---NHEKPADLGFDAANAYHTAAMDQPSIKWYVDGQ-LKHT 192
OY 123 EGGQVSNLJTGTCGLRFNWSSESA-AWGOFPDESKLP.LFOFINWVYVK 170
DB 193 ATTOIPAAFGK--IMMNLNMGTCVDDMLGSLYNGVN-PIYAHYDMKRYRK 238
```

```
RESULT 6
GUB_CLOTM STANDARD: PRT: 334 AA.
ID GUB_CLOTM
AC P29716; P37074;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase) (Laminarinase).
GN LICB OR LAMI.
OS Clostridium thermocellum.
OC Bacteria: Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
  Clostridium.
OX NCBI_TaxID=1515;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27405 / DSM 1237;
RX MEDLINE=92155194; PubMed=1740123;
RA Schimming S., Schwarz W.H., Staudenbauer W.L.;
RT "Structure of the Clostridium thermocellum gene licB and the encoded
  beta-1,3-1,4-glucanase. A catalytic region homologous to Bacillus
  lichenases joined to the reiterated domain of clostridial
  cellulases."
  Eur. J. Biochem. 204:13-19(1992).
RL
RN [2]
RP PRELIMINARY SEQUENCE FROM N.A.
RC STRAIN=P7;
RX MEDLINE=92095946; PubMed=1755832;
RA Zverlov V.V., Laptev D.A., Tishkov V.I., Velikodvorskaja G.A.;
RT "Nucleotide sequence of the Clostridium thermocellum laminarinase
  gene."
  Biochem. Biophys. Res. Commun. 181:507-512(1991).
RL
RN [3]
RP REVISIONS.
RA Zverlov V.V.;
RL Submitted (Aug-1993) to the EMBL/Genbank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RC STRAIN=P7;
RA Zverlov V.V., Velikodvorskaja G.A.;
RT "Cloning the Clostridium thermocellum thermostable laminarinase gene
  in Escherichia coli: the properties of the enzyme thus produced."
  Biotechnol. Lett. 12:811-816(1990).
RL
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
  in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- SUBUNIT: MAY FORM PART OF A MULTISUBUNIT COMPLEX (CELLULOSE).
CC -1- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS
  WELL AS IN OTHER C.THERMOCELLUM CELLULOSE ENZYMES. THIS DOMAIN
  MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
CC EMBL: X63355; CAA44959.1; -.
CC DR EMBL: X58392; CAA41281.1; -.
CC DR PIR: S23498; S23498.
CC DR PIR: JS0611; JS0611.
CC DR HSSP: P23904; 1A7K.
CC DR InterPro: IPR002105; Dockerin_1.
CC DR InterPro: IPR002048; EF-hand.
CC DR Pfam: PF00404; Dockerin_1; 2.
CC DR Pfam: PF00722; Glyco_hydro_16.
CC DR PRINTS: PR00737; GLHYDRASE16.
CC DR
```

```

DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
DR PROSITE: PS00018; EF_HAND; UNKNOWN; 2.
DR PROSITE: PS00448; CLOS_CELLOSOME_RP; 2.
KM Hydrolase; Glycosidase; Signal; Repeat.
FT SIGNAL 1 27
FT CHAIN 28 334
FT ACT_SITE 136 136 BETA-GLUCANASE.
FT ACT_SITE 140 140 NUCLEOPHILE (BY SIMILARITY).
FT DOMAIN 252 269 PROTON DONOR (BY SIMILARITY).
FT DOMAIN 273 331 PRO/THR-RICH (LINKER).
FT REPEAT 273 331 2 X 24 AA APPROXIMATE REPEATS.
FT REPEAT 308 331 1.
FT REPEAT 308 331 2.
FT CONFLICT 304 334
FT SEQUENCE 334 AA; 37897 MW; 0837564E9726F281 CRC64;
SO QUERY MATCH 16.5%; Score 237; DB 1; Length 334;
Best Local Similarity 30.9%; Pred. No. 1,3e-12;
Matches 80; Conservative 27; Mismatches 94; Indels 58; Gaps 13;
QY 5 KDSGAEIYTLLEVO-----YGFARMKMAAGTVSSMFLYONGSEIADGRPWVEVD 58
DB 82 REVGSGYPRKSGEYRKRFSGFYGYEVFRMKAANVGIVSSFFYTGPS--DNNPWDEID 138
QY 59 IEVLGNPGSFOSNIITGKAGAKTSEKHHAVPADQAFHTGTLEMTPTVYRWMTVDGQE 118
DB 139 IEFLGMDITTKVQGFNMWKNNGVG---NEYLHNLGFDASQDPHTGYGFEMRPDIIFYDGKK 195
QY 119 VKRTSGGVSNLTGTGCG-LRFNIMSSESA-AWVGQFDESKLPLEQFINWVKYV----- 169
DB 196 VVR-----GTRNIFVTGPKIMMNLMPGIGVDEWGRYD-GRTPLQAEYEVKYYPNGVPOD 250
QY 170 -----KYRP-----GGE-GGSDFT-----LDMWNPTFGSRKRGK 201
DB 251 NPLPTPTAPSTPTPNPLKGVNDGHNSSDYLFRKRYLLRLVIDRFPVGQS---VA 307
QY 202 DMTFDGNRYDLTDKNITYSR 220
DB 308 DVNRDG-RIDSTDLTMLKR 325
RESULT 7
GUB_BACLI STANDARD; PRT; 243 AA.
AC P27051:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
GN BGL.
OS Bacillus licheniformis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1402;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91224124; PubMed=2026156;
RA Lloberas J., Perez-Pons J.A., Querol E.;
RT "Molecular cloning, expression and nucleotide sequence of the
RT endo-beta-1,3-1,4-D-glucanase gene from Bacillus licheniformis.
RT Predictive structural analyses of the encoded polypeptide.";
RU Eur. J. Biochem. 197;337-343(1991).
RN [2]
RP REVISIONS.
RA Querol E.;
RL Submitted (JUL-1991) to the EMBL/Genbank/DBJ databases.
RN [3]
RP MUTAGENESIS.
RX MEDLINE=92362869; PubMed=1354172;
RA Planas A., Juncosa M., Lloberas J., Querol E.;
RT "Essential catalytic role of Glu134 in endo-beta-1,3-1,4-D-glucan 4-
RT glucanohydrolase from B. licheniformis as determined by site-directed
RT mutagenesis.";
```

```

RL FEBS Lett. 308:141-145(1992).
RN [4]
RP MUTAGENESIS.
RX MEDLINE=94237863; PubMed=8182059;
RA Juncosa M., Pons J., Dot T., Querol E., Planas A.;
RT "Identification of active site carboxylic residues in Bacillus
RT licheniformis 1,3-1,4-beta-D-glucan 4-glucanohydrolase by
RT site-directed mutagenesis.";
RL J. Biol. Chem. 269:14530-14535(1994).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX MEDLINE=96063718; PubMed=7589539;
RA Hahn M., Pons J., Planas A., Querol E., Heinemann U.;
RT "Crystal structure of Bacillus licheniformis 1,3-1,4-beta-D-glucan 4-
RT glucanohydrolase at 1.8-A resolution.";
RL FEBS Lett. 374:221-224(1995).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glucosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X57279; CAA0547.1; -
DR PIR; S1538G; S15388
DR PDB; 1GBG; 07-DEC-95.
DR InterPro; IPR000757; Glyco_hydro_16.
DR Pfam; PF00722; Glyco_hydro_16; 1.
DR PRINTS; PR00737; GLHYDRLASE16.
DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KM Hydrolase; Glycosidase; Signal; 3D-structure.
FT SIGNAL 1 27
FT CHAIN 28 243
FT ACT_SITE 134 134 BETA-GLUCANASE.
FT ACT_SITE 138 138 NUCLEOPHILE.
FT DISULFID 61 90
FT MUTAGEN 51 51 PROTON DONOR.
FT MUTAGEN 89 89 D->N: 30% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 92 92 D->N: 85% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 105 105 E->Q: 3% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 133 133 E->Q: 50% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 134 133 D->N: 15% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 136 134 E->Q: 0.2% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 136 136 D->N: 0.5% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 138 136 E->Q: COMPLETE LOSS OF ACTIVITY.
FT MUTAGEN 143 143 D->N: 65% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 160 160 E->Q: 15% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 168 168 D->N: 60% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 179 179 D->N: 80% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 190 190 D->N: 70% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 219 219 D->N: NO CHANGE IN ACTIVITY.
SQ SEQUENCE 243 AA; 27435 MW; 651189DAAD609A5 CRC64;
QUERY MATCH 15.9%; Score 229.5; DB 1; Length 243;
Best Local Similarity 33.7%; Pred. No. 3,8e-12;
Matches 57; Conservative 25; Mismatches 76; Indels 11; Gaps 6;
QY 3 SAKDFSGAEIYTLLEVOYQGFARMKMAAGTVSSMFLYONGSEIADGRPWVEVDIEVL 62
DB 84 SYNKPFGCGENRSYQVGYGLYEVNMRPAKNVGIVSSFFYTGP---DGPMDIEDIEFL 140
QY 63 GKNPGSFOSNIITGKAGAKTSEKHHAVPADQAFHTGTLEMTPTVYRWMTVDGQEVKRT 122
DB 141 GKDTTVQGFNMWKNNGVG---NHEKIVNLGFDANSTYTAFDQPSIKWYVDGO-LKHT 196
QY 123 EGQGVSNLTGTGGLRFLNIMSSESA-AWVGQFDESKLPLEQFINWVKYV 170
```

DB 197 ATTOIPOTPKG--IMNLMWAGVDEWLGSYN-GVTPPLSRSLHWRYTK 242

RESULT 8

GUB_PAEMA STANDARD; PRT; 237 AA.

ID GUB_PAEMA

AC P23904;

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Beta glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)

OS (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).

OC Paenibacillus macerans (Bacillus macerans).

OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.

OX NCBI_TaxID=44252;

RP (1)

RP SEQUENCE FROM N.A.

RX MEDLINE-91109712; PubMed-2274030;

RA Boriss R., Bueltner K., Maentsaele P.;

RT "Structure of the beta-1,3-1,4-glucanase gene of Bacillus macerans:

RT homologues to other beta-glucanases.";

RL MOL. Gen. Genet. 222:278-283(1990).

RN [2]

RP ACTIVE SITE.

RX MEDLINE-93094208; PubMed-1360982;

RA Hoej P.B., Condron R., Traeger J.C., McAuliffe J.C., Stone B.A.;

RT "Identification of glutamic acid 105 at the active site of Bacillus

RT amyloliquefaciens 1,3-1,4-beta-D-glucan 4-glucanohydrolase using

RT epoxide-based inhibitors.";

RL J. Biol. Chem. 267:25059-25066(1992).

RN [3]

RP MUTAGENESIS OF GLU-128.

RA Olsen O.;

RL Theists (1990), University of Aarhus, Denmark.

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).

RX MEDLINE-93281743; PubMed-8099449;

RA Keitel T., Simon O., Boriss R., Heinemann U.;

RT "Molecular and active-site structure of a Bacillus 1,3-1,4-beta-

RT glucanase";

RL Proc. Natl. Acad. Sci. U.S.A. 90:5287-5291(1993).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).

RX MEDLINE-96028129; PubMed-7588726;

RA Hahn M., Keitel T., Heinemann U.;

RT "Crystal and molecular structure at 0.16-nm resolution of the hybrid

RT Bacillus endo-1,3-1,4-beta-D-glucan 4-glucanohydrolase H(A16-M).";

RL Eur. J. Biochem. 232:849-858(1995).

CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages

CC in beta-D-glucans containing 1,3- and 1,4-bonds.

CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE

CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.

CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

CC EMBL; X55959; CAA39426.1; -

DR PIR: S11927; S11927.

DR PDB: 2AVH; 31-MAR-95.

DR PDB: 1BYH; 31-OCT-93.

DR PDB: ICPM; 22-JUN-94.

DR PDB: ICPM; 22-JUN-94.

DR PDB: IMAC; 27-FEB-95.

DR PDB: IATK; 06-MAY-98.

DR PDB: IATO; 06-MAY-98.

DR InterPro: IPR000757; Glyco_hydro_16.

DR Pfam: PF00722; Glyco_hydro_16; 1.

DR PRINTS: PR00737; GLHYDRLASE16.

DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.

KW Hydrolase; Glycosidase; Signal: 3d-structure.

FT SIGNAL 1 23

FT CHAIN 24 237

FT ACT_SITE 128 128 BETA-GLUCANASE.

FT ACT_SITE 132 132 NUCLEOPHILE (BY SIMILARITY).

FT DISULFID 55 84 PROTON DONOR (BY SIMILARITY).

FT MUTAGEN 128 128 E->D,N,A,L,P,R,H,C,S,Y: LOSS OF ACTIVITY.

FT STRAND 26 26

FT STRAND 29 31

FT STRAND 41 43

FT STRAND 51 52

FT STRAND 53 53

FT STRAND 56 57

FT STRAND 59 61

FT HELIX 62 64

FT STRAND 66 67

FT STRAND 70 78

FT STRAND 79 80

FT STRAND 81 89

FT STRAND 93 93

FT STRAND 96 103

FT STRAND 108 109

FT STRAND 110 118

FT STRAND 120 122

FT HELIX 123 123

FT STRAND 127 134

FT STRAND 135 136

FT STRAND 137 147

FT TURN 148 149

FT STRAND 150 150

FT STRAND 155 158

FT STRAND 163 164

FT STRAND 168 174

FT STRAND 179 183

FT TURN 184 185

FT STRAND 186 191

FT STRAND 200 209

FT STRAND 213 216

FT STRAND 225 236

SQ SEQUENCE 237 AA; 26589 MW; 436EABCDFFC87781 CRC64;

Query Match 15.7%; Score 226; DB 1; Length 237;

Best local Similarity 35.1%; Pred. No. 7,2e-12;

Matches 59; Conservative 23; Mismatches 70; Indels 16; Gaps 8;

QY 7 FSCAELYTLEEVQYQKFEARMKMAASGVSSMFLYQNGSEIADGRPWVEDIEVLGKNP 66

DB 82 FDCAEFRSTNIYGYGLYEVMKPAKNTGIVSSFFTYTGP--AHGTQMEIDIEFLGKDT 138

QY 67 GSFQSNIIITGKKAQOKTSKHNHVAADQAFHTYGLFWTPRYVRYNTVDGQEVKRTGGQ 126

DB 139 TKVQFNYVTNGVGH--RKVISTLGFDAKSGFHTYAFDQPYIMWYDG-VLKHT---A 191

QY 127 VSNLTGTQG-LRFNLWSSESA-AMWGFDESKLPLFOFINWKKVYKYT 172

DB 192 TANIPSTPKIMNLMNGVCDWDLGISTYNGAN-PLXAEYDWY---KYT 235

RESULT 9

GUB_PAEPD STANDARD; PRT; 238 AA.

ID GUB_PAEPD

AC P45797;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)

OS (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).

GN GLUB.

OC Paenibacillus polymyxa (Bacillus polymyxa).

OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.

OX NCBI_TaxID=1406;

```

RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 842;
RX MEDLINE=92041687; PubMed=1938968;
RA Gosalbes M.J., Perez-Gonzalez J.A., Gonzalez R., Navarro A.;
RT "Two beta-glycanase genes are clustered in Bacillus polymyxa:
RT molecular cloning, expression, and sequence analysis of genes
RT encoding a xylanase and an endo-beta-(1,3)-(4,4)-glucanase.";
RL J. Bacteriol. 173:7705-7710(1991).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X57094; CAA40379.1; -.
DR HSSP: P23904; ICPN.
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro.16; 1.
DR PRINTS: PR00737; GLHYDRLASE16.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Hydrolyase; Glycosidase; Signal.
FT SIGNAL 1 26
FT CHAIN 27 238
FT ACT_SITE 129 133
FT ACT_SITE 133 133
FT DISULFID 56 85
FT SEQUENCE 238 AA; 26919 MW; C0CF7B4EASD40E8C CRC64;

Query Match 15.6%; Score 225; DB 1; Length 238;
Best Local Similarity 34.5%; Pred. No. 8, 8e-12;
Matches 58; Conservative 23; Mismatches 71; Indels 16; Gaps 8;

QY 7 FSCGAEIYTLDEVOYGFKEARMKMAASGVSSMFLYONGSEIADGRPWVEVDIEVLGKNP 66
DB 83 FDCEGRSTNNYGLIYEVSMKPAKNTGIIVSEFTYGPSSH---GTQWDEIDIEFLGKDT 139
QY 67 GSFSQNIITKAGAKQKSEKHHAVSPADAQAFHTYGLEMPNRYRWTVDDQEVKTKSGGQ 126
DB 140 TKYQFNYYTNGVGCH--EKIINLGFDASTSFHYAFDQWOPGYTKWYVDG-VLKHT--A 192

QY 127 VSNLTGTGOG-LRFLWSSESA-AWVGQFDESKLPLFOFIWVKVYKYT 172
DB 193 TTNIPSPGKIMNMLWNGTGVDSWLSGYNGAN-PLVAYEYDWV---KKT 236

RESULT 10
EXOK_RHIME STANDARD; PRT: 269 AA.
AC P33693;
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Endo-1,3-1,4-beta-glycanase exok precursor (EC 3.2.1.-) (Succlnogl1can
DE biosynthesis protein exok).
GN EXOK OR RB1080 OR SMB20955.
OS Rhizobium melioli (Sinorhizobium melioli).
OC Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN RP
SEQUENCE FROM N.A.
STRAIN=1021.
RX MEDLINE=94042869; PubMed=8226645;

```

```

RA Glucksman M.A., Reuber T.L., Walker G.C.;
RT "Family of glycosyl transferases needed for the synthesis of
RT succinoglycan by Rhizobium melioli."
RL J. Bacteriol. 175:7033-7044(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RX MEDLINE=93241147; PubMed=8479421;
RA Backer A., Kleickmann A., Arnold W., Puehler A.;
RT "Analysis of the Rhizobium melioli exoH/exoK/exoL fragment: ExoK
RT shows homology to excreted endo-beta-1,3-1,4-glucanases and ExoH
RT resembles membrane proteins."
RL Mol. Gen. Genet. 238:145-154(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=1021.
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorheiter F.J., Hernandez-Lucas I., Backer A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium melioli."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=98226741; PubMed=9560202;
RA York G.M., Walker G.C.;
RT "The Rhizobium melioli ExoK and ExoH glycanases specifically
RT depolymerize nascent succinoglycan chains."
RL Proc. Natl. Acad. Sci. U.S.A. 95:4912-4917(1998).
CC -1- FUNCTION: CLEAVES HIGH MOLECULAR WEIGHT SUCCINOGLYCAN TO YIELD LMW
CC SUCCINOGLYCAN. DYNAMICALLY REGULATE THE MOLECULAR WEIGHT
CC DISTRIBUTION OF SUCCINOGLYCAN BY CLEAVING NASCENT SUCCINOGLYCAN
CC ONLY DURING A LIMITED PERIOD AFTER ITS SYNTHESIS, PERHAPS BEFORE
CC IT UNDERGOES A TIME-DEPENDENT CHANGE IN ITS CONFORMATION OR
CC AGGREGATION STATE.
CC -1- PATHWAY: Exopolysaccharide biosynthesis.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L20758; AAA16048.1; -.
DR EMBL: Z17219; CAA73927.1; -.
DR EMBL: A1603645; CAC49480.1; -.
DR HSSP: P23904; IAKJ.
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro.16; 1.
DR PRINTS: PR00737; GLHYDRLASE16.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Exopolysaccharide synthesis; Glycosidase; Hydrolase; Signal; Plasmid;
KW Complete proteome.
FT SIGNAL 1 27
FT CHAIN 28 269
FT CONFLICT 93 102
FT SEQUENCE 269 AA; 30083 MW; 073CC7ED65EF2611 CRC64;

Query Match 13.7%; Score 197; DB 1; Length 269;
Best Local Similarity 32.5%; Pred. No. 2, 2e-09;
Matches 50; Conservative 31; Mismatches 61; Indels 12; Gaps 6;

QY 2 VSAQDFGAEIYTLDEVOYGFKEARMKMAASGVSSMFLYONGSEIADGRPWVEVDIEV 61
DB 87 VKERNFACGEIQTFKRGYGTIERKRAADSGSLNSAFFTYIGP---ADKKPDEIDIEV 143
QY 62 LGRNPGSFQSN-IITKAGAKQKSEKHHAVSPADAQAFHTYGLEMPNRYRWTVDDQGEVR 120

```

Db 144 LGKNTAKVQJNOYVSAKGNFLAD---VPGCANOGFNDYAFWEXNRIRYYVNGELVH 199
 Oy 121 K-TEGGGVSNLTGTGRLFNLMSESA--WVGOF 152
 Db 200 EVTIDPAKIP--VNAOKIFFSLMGITDILTDMWGT 231

RESULT 11
 YG46.YEAST STANDARD: PRT: 507 AA.

AC P53301;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypochemical 52.8 kDa protein in BUB1-HIP1 Intergeric region.
 GN YGR189C OR G7553.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=5288C;
 RX MEDLINE=97279231; PubMed=9133739;
 RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,
 RA Nombela C.;
 RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
 of Saccharomyces cerevisiae chromosome VII.",
 RL Yeast 13:357-363(1997).
 CC -1- SIMILARITY: SOME, TO YEAST UTR2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: 272974; CAA97215.1;
 DR EMBL: X99074; CAA67525.1;
 DR HSSP: P23904; 1AUK.
 DR SGI: S0003421; CRH1.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 KW Hypochemical protein.
 FT DOMAIN 63 66 POLY-SER.
 FT DOMAIN 301 310 POLY-SER.
 FT DOMAIN 345 357 POLY-SER.
 FT DOMAIN 387 391 POLY-SER.
 FT DOMAIN 467 470 POLY-SER.
 SQ SEQUENCE 507 AA: 52757 MW: 707861F57AE942C CRC64;

Query Match 12.78; Score 182.5; DB 1; Length 507;

Best Local Similarity 26.78; Pred. No. 7.4e-08; Indels 31; Gaps 8;

Matches 58; Conservative 37; Mismatches 91;

Oy 1 MVSADPSGAELTYLLEVOYGFEPARKMAAGTSMFLYONGSFIAGRPVVEVDIE 60
 Db 87 MTLAKRRDPSLSNFRIMYKLEVLKANGTIVSFFLQSDDD------EIDIE 138
 Oy 61 VLKRNPSFOSNITTKAGAKOXTSEKHAVSPADDAFHTYGTLEMPNYVMTVDGQEV 120
 Db 139 WVGDMNTQFOSNFFSKGDDTTYDRGEFHGYDPTDK-FHHYITLDMAMDKTTWLDSGSR 197
 Oy 121 KTEGGGVSNLTGTG-----LRFNLMS--ESAA---WV-QGFDESKLPLDFQFINW 165
 Db 198 -----VLSNTSSEGYQPSQPMYLMGIMAGGDDPNAAGTIEMAGETNNYNDAPFTWIEK 251

Oy 166 VKYKYTTPGGEGSDFTLDMTNTFDTFGSKRGKD 202
 Db 252 VVITDYSTGKRYVGDQSGM-ESIADGGSIGYGRD 287

RESULT 12

MERS_ARATH STANDARD: PRT: 269 AA.

ID MERS_ARATH
 AC P24806; Q39148; Q41904; Q64956;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE MER1-5 protein precursor (Endo-xyloglucan transferase) (Xyloglucan
 DE endo-1,4-beta-D-glucanase).
 GN MER1-5 OR MER15B OR SEN4 OR AT4G30270 OR F9N11.120.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93005704; PubMed=1840916;
 RA Medford J.I., Elmer J.S., Klee H.J.;
 RT "Molecular cloning and characterization of genes expressed in shoot
 RT apical meristems";
 RL Plant Cell 3:359-370(1991).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Kaminal T., Tomita E., Nishitani K.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Arrowsmith D.A.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.

RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Scheller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Ertlan K.-D., Terry N.,
 RA Harris B., Ansoorge W., Brandt P., Griwell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermair B., Maché R., Mueller M.,
 RA Kreis M., Delzeny M., Puigdomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohnel S., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Blahm L., Robben J.,
 RA Van der Schueren J., Glymoprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weltenegeger T., Botne G., Rampsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirxse W.,
 RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Koeltter P.,
 RA Bernier S., Hempel S., Feldpausch M., Lambert S., Van den Daele H.,
 RA de Keyser A., Buysheert C., Gielen J., Villarejo R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Brey-Alen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
 RA Pettelt A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Granderath K., Danner D., Herzi A.,
 RA Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Cheford F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bialek C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker C.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schütz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,

RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
 RA Antoniou B., Zidanovic A., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Val D., Shekher M., Matero A., Shah R.,
 RA Swaby T.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
 RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
 RA Chen E., Maria M., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana.";
 RT Nature 402:769-777(1999).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SSP consortium (Salk/Stanford/PGEC).";
 RT Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 1-132 FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=Green silique;
 RA Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delseny M.;
 RL Submitted (OCT-1992) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 1-120 FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=Leaf;
 RX MEDLINE=98278374; PubMed=9617812;
 RA Park J.-H., Oh S.A., Kim Y.H., Woo H.R., Nam H.G.;
 RT "Differential expression of senescence-associated mRNAs during leaf
 RT senescence induced by different senescence-inducing factors in
 RT Arabidopsis.";
 RT Plant Mol. Biol. 37:445-454(1998).
 CC [1]
 CC -1- FUNCTION: Involved in cell wall reconstruction.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN SHOOT APICAL MERISTEMS, ALSO
 CC FOUND IN SEEDLINGS AND MERISTEMS.
 CC [2]
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
 CC in positions 158, 178, 183, 189, 190, 194 and 199.
 CC -1- CAUTION: Ref.6 sequence differs from that shown due to frameshifts
 CC in positions 93 and 104.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M63166; AAA32828.1; ALT_FRAME.
 DR EMBL: D63508; BAA09783.1; -.
 DR EMBL: AL109796; CAB52471.1; -.
 DR EMBL: AL161576; CAB81020.1; -.
 DR EMBL: AY035156; AAK59660.1; -.
 DR EMBL: AY063027; AAL34201.1; -.
 DR EMBL: Z17602; CAA79012.1; ALT_FRAME.
 DR EMBL: AF035384; AAC39467.1; -.
 DR PIR: J01022; J01022.
 DR HSSP: P23904; IAUJ.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 KW Glycosidase; Hydrolase; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 269 MERI-5 PROTEIN.
 FT CONFLICT 93 94 A -> G (IN REF. 1).
 FT CONFLICT 184 184 A -> G (IN REF. 7).
 SQ SEQUENCE 269 AA; 30755 MW; 648F042BC7ADE86 CRC64;

Query Match 10.5%; Score 150.5; DB 1; Length 269;
 Best Local Similarity 26.3%; Pred. No. 1.6e-05;
 Matches 47; Conservative 34; Mismatches 65; Indels 33; Gaps 8;

QY 8 SGAEYLTLEVOYGFKFEARRKMAA--ASGTIVSSMPLYONGSFIAIDGRPWVEVDIEVLGK- 64

DB 54 SGGSGFOSKTEYLFGRKIDMQIKLVPGNSAGVTTFYLSKSEGS-----TWDEIDFEFLGNM 107
 QY 65 --NPGSFQSNITIGKACAGCTSEKHNAVSPADAQAFITYLEMTPTNVRFTVVGQEVKRT 122
 DB 108 SGPYLTHTFTVWYQSGXG-DKEQGFHLMFDPETAN--FHTYSILNNPQRIILTVDPTPIRF 164
 QY 123 EGGGVSNLTCTGGLRF-----NLWSESAWVG--GFDESKLPFLQFINWVKV 168
 DB 165 KNYE-----SLGYLFFPKNPMRRYASLWNAQDMATRTGVLGKTDWSKAPMAYSTRINKI 217
 RESULT 13
 UTR2 YEAST STANDARD; PRT; 347 AA.
 AC P32623;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE UTR2 protein (unknown transcript 2 protein).
 GN UTR2 OR YEL040W OR SYGP-ORF18.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,
 RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
 RL Submitted (FEB-1993) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunkle-Smith S.,
 RA Hyman R., Kayser A., Komp C., Laskari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Narmath A., Norgren R., Oefner P., Oh C.,
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B-6441;
 RX MEDLINE=94016558; PubMed=8411151;
 RA Melnick L., Sherman F.;
 RT "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,
 RT of Saccharomyces cerevisiae share a common ancestry.";
 RL J. Mol. Biol. 233:372-388(1993).
 CC [1]
 CC -1- SIMILARITY: SOME, TO YEAST YGR189C.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U18779; AAB65002.1; ALT_INIT.
 DR EMBL: L22173; AAA34941.1; -.
 DR EMBL: S65964; AAD13975.1; -.
 DR EMBL: S66130; AAB28444.1; -.
 DR PIR: S30839; S30839.
 DR HSSP: P23904; IAUJ.
 DR SGD: S0000766; UTR2.
 DR InterPro: IPR000757; Glyco_hydro_16; 1.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 FT DOMAIN 234 322 SER-RICH.
 FT DOMAIN 269 283 POLY-SER.
 FT CONFLICT 10 10 L -> V (IN REF. 3).
 FT CONFLICT 171 171 A -> R (IN REF. 3).
 FT CONFLICT 234 234 S -> C (IN REF. 3).

Thu Jan 9 12:09:16 2003

us-09-654-652a-2.rsp

Page 11

Db 99 SQGPTHDEIDPEFLGNLSGDPYLTHNTIEFLQGKG-NNEQGFYLMFPDPTRN--PHTYSIIM 155

Qy 106 TPNNVYVQOGEVREKREGGGVSNL--TGTQGRF--NLSSSESAMVG--QFDESKIP 158

Db 156 KPGHIEFLVDNTPFRVRKNAPEPLGVPPPKQNPARKIYSSLNADMDATRGSLVATDWSKAP 215

Qy 159 LFOFINWYKYKYKT 172

Db 216 FTATYRNKAEIES 229

```
Search completed: January 9, 2003, 12:16:48
Job time : 8.79861 secs
```



```

DB      93 LENDQAPPYKAGELRTNDFQYGLFEVSMKPAKVEGVSSFFTY-TGEMDMDGPMPEI 151
OY      81 DIEVGKPNQSGFQSNITITGKAGAOCTSEKHNAVSPAADOAFHTYCLEMTPNVVRVTVDGQ 140
DB      152 DIEFLGKQTRTQFNYFTNGVG---NEFYDLGFDASESFNTYAFWEHREDSITWYNGE 208
OY      141 EVRKTEGGOVSNLTCT-QGLRFNLWSSSA-AWVGQFDESKLPFQFINWVVKYKTP-- 196
DB      209 AVHTA-----TENIPOTPOKIMNMLMPGVGVDENITVFEGDNTPLHADYEMV---RYTPLE 261
OY      197 --GGGEGSSDFTLDMT-----DNEDFPDGSRMG-KGDWTFDG 230
DB      262 ALDEESGDNDNEEVEEVEEPEADEEVSVPISALYETFTENTNEDIWSIAHGWT-NG 320
OY      231 NRVDLT--DKNIYSDGHLIALATRKQGESFNGQVPRDEPA 270
DB      321 QMFNTWYDSQITTFNSGIMRAIDK-----DDEEA 351

```

RESULT 2

```

O9APD8 PRELIMINARY: PRT: 256 AA.
AC      01-JUN-2001 (TREMBlrel. 17, Created)
DT      01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DR      01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE      Endo-beta-1,3-1,4-glucanase.
CN      BGLBCL.
OS      Bacillus circulans.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX      Bacillaceae; Bacillus.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-ATCC21367.
RA      Lee D.-S., Kim J.-Y., Kim H.-B.;
RL      Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF306531; AAC53947.1; -.
DR      HSSP: P23904; IATK.
DR      InterPro: IPR000757; Glyco_hydro_16.
DR      Pfam: PF00722; Glyco_hydro_16.
DR      PRINTS: PR00737; GLHYDRASE16.
DR      PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
SQ      SEQUENCE 256 AA; 27604 MW; 989E50C09F64531 CRC64;

```

```

Query Match      14.0%; Score 254.5; DB 2; Length 256;
Best Local Similarity 34.6%; Pred. No. 1.3e-09;
Matches 62; Conservative 29; Mismatches 71; Indels 17; Gaps 7;

```

```

OY      23 TNVSAKDPGSAELTYLEVOYGKFEARKMAAAGTVSSMFLYONGSEIADGRPWVEVDI 82
DB      88 SNGSGKPRASAEVATQYGYGRVBARIKAKGTGLVTSLTFT---SCAACTSDNDEIDI 144
OY      83 EVLKNPGSFQSNITITGKAGAOCTSEKHNAVSPAADOA-FHTYCLEMTPNVVRVTVDGQ 140
DB      145 EILCKDTTKMETNYFTNGVCGHST-----VIDLGFDAISLDPHDAFAEWSPTSIKWYVGR 199
OY      141 EVRTEGGOVSNLTCTGQ-LRFNLMSSSA--WVGQDESKLPFQFINWVVKYKTP 196
DB      200 LVH-TETGSRGPLPTSPQIYVNMISGAGPAEIMTGTGTPHPIRAYDVI---KFTP 254

```

RESULT 3

```

O45648 PRELIMINARY: PRT: 276 AA.
AC      01-NOV-1996 (TREMBlrel. 01, Created)
DT      01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DR      01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE      Lichenase precursor.
CN      BGL1.
OS      Bacillus sp.

```

```

OC      Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX      NCBI_TaxID=1409;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-137;
RX      MEDLINE-94286605; PubMed=7517127;
RA      Taberner C., Coll P.M., Fernandez-Abalos J.M., Perez P.,
RN      Santamaria R.I.;
RT      Cloning and DNA sequencing of bgaA, a gene encoding an endo-beta-1,3-
RT      1,4-glucanase, from an alkalophilic Bacillus strain (N137).";
RL      Appl. Environ. Microbiol. 60:1213-1220(1994).
DR      EMBL: 212151; CAAT8135.1; -.
DR      HSSP: P23904; IATK.
DR      InterPro: IPR000757; Glyco_hydro_16.
DR      Pfam: PF00722; Glyco_hydro_16; 1.
DR      PRINTS: PR00737; GLHYDRASE16.
DR      PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
FT      SIGNAL.
FT      CHAIN.
SQ      SEQUENCE 276 AA; 31770 MW; D047F8A34CA9EBE2 CRC64;

```

```

Query Match      13.6%; Score 247.5; DB 2; Length 276;
Best Local Similarity 36.1%; Pred. No. 4.3e-09;
Matches 61; Conservative 20; Mismatches 75; Indels 13; Gaps 6;

```

```

OY      30 FSGAELTYLEVOYGKFEARKMAAAGTVSSMFLYONGSEIADGRPWVEVDIYGLKNP 89
DB      96 YKAGELRTNDFHYHGLFEVSMKPAKSTGTSSFFTY-TGPDWENDPDWDEIDIEFLGKDT 154
OY      90 GSFQSNITITGKAGAOCTSEKHNAVSPAADOAFHTYCLEMTPNVVRVTVDGQVKTGCGQ 149
DB      155 TKIQFNFTNGVG---NEHYHGLFDAADDFNTYAEWRPESIRWVNGSELVITA---- 207
OY      150 VSNLTGT-QGLRFNLWSSSA-AWVGQFDESKLPFQFINWVVKYKTP 196
DB      208 TENIPOTPOKIMNMLMPGIGVDGWTGRFNGEDTVPVYQYDVI---KFTP 253

```

RESULT 4

```

O93GE8 PRELIMINARY: PRT: 214 AA.
AC      093GE8;
DT      01-DEC-2001 (TREMBlrel. 19, Created)
DT      01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT      01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE      Endo 1-3,1-4-beta-glucanase (fragment).
OS      uncultured bacterium.
OC      Bacteria; environmental samples.
OX      NCBI_TaxID=77133;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Rodriguez V., Mellado R.P.;
RT      "Novel lichenases from soil.";
RL      Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF254957; AAK50610.2; -.
DR      InterPro: IPR000757; Glyco_hydro_16.
DR      Pfam: PF00722; Glyco_hydro_16; 1.
DR      PROSITE: PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT      NON_TER.
SQ      SEQUENCE 214 AA; 24159 MW; 3DB186D3B9D991CF CRC64;

```

```

Query Match      13.5%; Score 246.5; DB 2; Length 214;
Best Local Similarity 35.5%; Pred. No. 3.6e-09;
Matches 61; Conservative 25; Mismatches 75; Indels 11; Gaps 6;

```

```

OY      23 TNVSAKDPGSAELTYLEVOYGKFEARKMAAAGTVSSMFLYONGSEIADGRPWVEVDI 82
DB      52 TSPSYNKFDCGNSVOTTYGYGRVBARIKAKGTGLVTSLTFT---GTPWDEIDI 108
OY      83 EVLKNPGSFQSNITITGKAGAOCTSEKHNAVSPAADOAFHTYCLEMTPNVVRVTVDGQEV 142

```

```
Dh 109 EFLCKDVTYKVOFNYYTNGAG---NHEKRLADLGFDAANAHTYAFDWPQNSIKMYVDGQ-L 164
Qy 143 RKTGCGOVSNLTGTGGLRFLNMSSESA-AWVGOFDESKLPLFOFINNVKYYK 193
Dh 165 KHTATTTQIPAPGK--IMMNLNMGTVGVDMLGSYNGVN-PLYAHYDWRVRYTK 213

RESULT 5
093GE7 PRELIMINARY; PRT; 214 AA.
AC 093GE7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Endo 1-3,1-4-beta-glucanase (Fragment).
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez V., Melado R.P.;
RT "Novel lichenases from soil.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF254959; AAK50612.2;
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR PRINTS: PR00737; GLYCOLASE16.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 214 AA; 24175 MW; 3DB192D2F6B4CAA3 CRC64;

Query Match 13.5%; Score 245.5; DB 2; Length 214;
Best Local Similarity 35.5%; Pred. No. 4.2e-09;
Matches 61; Conservative 25; Mismatches 75; Indels 11; Gaps 6;

Qy 23 TNVSAKDFSGAELTYLEVOYGFKEARKMAAAGTVSMFLYONGSEIADGRPWVEVDI 82
Dh 52 TSPSYNKFDCGENSVOTYGYGLYEVRMKPAKNTGIVSSFTTYGPTGPE---GTPWDEIDI 108
Qy 83 EVLGKNGSFQSNITITKAGAKTSEKHHAVSPADAQFHTYGLGEMTPNRYRWTVDDQEV 142
Dh 109 EFLCKDVTYKVOFNYYTNGAG---NHEKRLADLGFDAANAHTYAFDWPQNSIKMYVDGQ-L 164
Qy 143 RKTGCGOVSNLTGTGGLRFLNMSSESA-AWVGOFDESKLPLFOFINNVKYYK 193
Dh 165 KHTATTTQIPAPGK--IMMNLNMGTVGVDMLGSYNGVN-PLYAHYDWRVRYTK 213

RESULT 6
09S310 PRELIMINARY; PRT; 802 AA.
AC 09S310;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Family 11 xylanase /family 16 beta (1,3-1,4) glucanase.
GN XYND.
OS Ruminooccus flavofaciens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Lachnospiraceae; Ruminooccus.
OX NCBI_TaxID=1265;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=17;
RC Aurilia V., Martin J.C., Munro C.A., Mercer D.K., Flint H.U.;
RT "Organisation and strain distribution of genes responsible for the
RT utilisation of xylans by the rumen cellulolytic bacterium Ruminooccus
RT flavofaciens 17.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ132472; CAB51934.1;
DR HSP; P23904; IAKJ.
DR InterPro: IPR003305; CBM_Cenc.
DR InterPro: IPR001137; GH_11.
```

```
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF002018; CBM 4.9; 2.
DR Pfam: PF00457; Glyco_hydro_11; 2.
DR Pfam: PF00722; Glyco_hydro_16; 2.
DR PRINTS: PR00911; GLAYDLASE11.
DR PRINTS: PR00737; GLYDLASE11.
DR PROSITE: PS00776; GLYCOSYL_HYDROL_F11.1; 2.
DR PROSITE: PS00777; GLYCOSYL_HYDROL_F11.2; 2.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 2.
FT CHAIN 2 F02
FT FAMILY 11 XYLANASE /FAMILY 16 BETA
SQ SEQUENCE 802 AA; 89019 MW; F36BC68805FC5274 CRC64;

Query Match 13.3%; Score 242.5; DB 2; Length 802;
Best Local Similarity 34.9%; Pred. No. 3.7e-08;
Matches 58; Conservative 21; Mismatches 74; Indels 13; Gaps 6;

Qy 30 FSGAELTYLEVOYGFKEARKMAAAGTVSMFLYONGSEIADGRPWVEVDIELGKNP 89
Dh 634 YSGGEFRTNNRYHYGYECSSQAKKNDGVSSFTTYGTPS---DGNFWEIDIEILGKNT 690
Qy 90 GSFSQNIITGKAGAKTSEKHHAVSPADAQFHTYGLGEMTPNRYRWTVDDQEVKTEGQ 149
Dh 691 TVQGFNYYTNGQKH---EKLYDGFDSSEAYHTYGFDPWPNTYAMVYDGREVYRA---- 743
Qy 150 VSNLTGTG--LRFLNMSSESA-AWVGOFDESKLPLFOFINNVKYYK 193
Dh 744 TQDLPKTPGKIMNAMPGLTVDDMLKAFN-GRTPLTAHYQWVYTK 788

RESULT 7
093GE6 PRELIMINARY; PRT; 214 AA.
AC 093GE6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Endo 1-3,1-4-beta-glucanase (Fragment).
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez V., Melado R.P.;
RT "Novel lichenases from soil.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF254961; AAK50614.2;
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 214 AA; 24131 MW; 1E9DBBFA300EB45 CRC64;

Query Match 13.1%; Score 239.5; DB 2; Length 214;
Best Local Similarity 35.5%; Pred. No. 1.1e-08;
Matches 61; Conservative 24; Mismatches 76; Indels 11; Gaps 6;

Qy 23 TNVSAKDFSGAELTYLEVOYGFKEARKMAAAGTVSMFLYONGSEIADGRPWVEVDI 82
Dh 52 TSPSYNKFDCGENSVOTYGYGLYEVRMKPAKNTGIVSSFTTYGPTGPE---GTPWDEIDI 108
Qy 83 EVLGKNGSFQSNITITKAGAKTSEKHHAVSPADAQFHTYGLGEMTPNRYRWTVDDQEV 142
Dh 109 EFLCKDVTYKVOFNYYTNGAG---NHEKRLADLGFDAANAHTYAFDWPQNSIKMYVDGQ-L 164
Qy 143 RKTGCGOVSNLTGTGGLRFLNMSSESA-AWVGOFDESKLPLFOFINNVKYYK 193
Dh 165 KHTATTTQIPAPGK--IMMNLNMGTVGVDMLGSYNGVN-PLYAHYDWRVRYTK 213

RESULT 8
04S691 PRELIMINARY; PRT; 239 AA.
ID 04S691
```

AC Q45691: 01-NOV-1996 (Tremblrel. 01, last sequence update)
 DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE Endo-beta-1,3-1,4-glucanase.
 GN BEG1.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCIB 8565;
 RA van Rensburg P., van Zyl W.H., Pretorius I.S.;
 RT "Over-expression of the Saccharomyces cerevisiae exo-beta-1,3-
 glucanase gene together with the Butyrivibrio fibrisolvens endo-beta-
 1,4-glucanase gene and the Bacillus subtilis endo-beta-1,3-1,4-
 glucanase gene in Saccharomyces cerevisiae."
 RT Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RL EMBL: U60830; ABO5759.1;
 DR HSSP: P27051; IGBG.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR PRINTS: PR00737; GLHYDRLASE16.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 SQ SEQUENCE 239 AA; 26969 MW; CA219760D13F878A CRC64;

Query Match 13.1%; Score 238.5; DB 2; Length 239;
 Best Local Similarity 33.7%; Pred. No. 1.4e-08;

Matches 59; Conservative 27; Mismatches 72; Indels 17; Gaps 7;

OY 23 TNVSAKDSGAEELTYLEVOYGFKEARKMAASGTVSMPLYONGSEIADGRPVEVDI 82
 DB 77 TSPSNKRDCCGNSVQTYGLEYVRKPKAKNGIVSSFTTYGPTG---CTPWDEIDI 133
 OY 83 EVLGKNGPSFQSNITITKAGAKTSEKHAVSPA---ADQFHTYGLEMTPNVYRWTVDG 139
 DB 134 EFLGKDTTKVQFNNTYNGA-----ENHEKLADLGFDAANAYHTYAFDMQNSIKWYDW 187
 OY 140 QEVKRTGGOVSNLTGTGGLRFNLWSSESA-WVCGFDESKLPLEQFINWKKYK 193
 DB 188 Q-LKHTATTQIPTPAGK--IMNMLNGTGVSDWLGSYNGVN-PLYAEYDWMVRYTK 238

RESULT 9

Q8RMP0 PRELIMINARY; PRT; 242 AA.

AC Q8RMP0: 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE Beta-1,3-1,4-glucanase.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sun J., Li W., Gu S., Xu Z., Zhao H., Xiao J.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF490978; AAM08358.1;
 SQ SEQUENCE 242 AA; 27238 MW; 285FED3FE76AE69A CRC64;

Query Match 13.0%; Score 236.5; DB 2; Length 242;
 Best Local Similarity 34.9%; Pred. No. 2e-08;

Matches 60; Conservative 25; Mismatches 76; Indels 11; Gaps 6;

OY 23 TNVSAKDSGAEELTYLEVOYGFKEARKMAASGTVSMPLYONGSEIADGRPVEVDI 82
 DB 80 TSPSNKRDCCGNSVQTYGLEYVRKPKAKNGIVSSFTTYGPTG---DGTWDEIDI 136
 OY 83 EVLGKNGPSFQSNITITKAGAKTSEKHAVSPAADQFHTYGLEMTPNVYRWTVDGQEV 142

DB 137 EFLGKDTTKVQFNNTYNGA---NHEKIVDGFDAANAYHTYAFDMQNSIKWYDQ-L 192
 OY 143 KRTGGOVSNLTGTGGLRFNLWSSESA-AWVGQDESKLPLEQFINWKKYK 193
 DB 193 KHTATNQIPTPGK--IMNMLNGTGVSDWLGSYNGVN-PLYAEYDWMVRYTK 241

RESULT 10

Q9FDC9 PRELIMINARY; PRT; 212 AA.

AC Q9FDC9: 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE Endo-1,3-1,4-beta-glucanase (fragment).
 OS Paenibacillus polymyxa (Bacillus polymyxa).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Paenibacillaceae; Paenibacillus.
 OX NCBI_TaxID=1406;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yao W., Wang Y., Song W., Yang K., Su Z.;
 RT "Gene cloning of an antifungal protein."
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF284449; AAC02415.1;
 DR HSSP: P23904; IAKJ.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR PRINTS: PR00737; GLHYDRLASE16.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 FT NON-TER 1 1
 FT NON-TER 212 212
 SQ SEQUENCE 212 AA; 24127 MW; 7E707E7EBE7A440 CRC64;

Query Match 12.7%; Score 232.5; DB 2; Length 212;
 Best Local Similarity 34.8%; Pred. No. 3.1e-08;

Matches 62; Conservative 26; Mismatches 73; Indels 17; Gaps 9;

OY 20 ALTNVSAKDSGAEELTYLEVOYGFKEARKMAASGTVSMPLYONGSEIADGRPVE 79
 DB 48 SLTSSAIK-PDGGVYSKNTYRGLEYVRKPKAKNGIVSSFTTYGP---ANGTQWDE 103
 OY 80 VDIKNGPSFQSNITITKAGAKTSEKHAVSPAADQFHTYGLEMTPNVYRWTVDG 139
 DB 104 IDIEFLGKDTTKVQFNNTYNGIGH---EKVVDLGFASSGFHTYARDMQPGYIKWYDG 160
 OY 140 QEVKRTGGOVSNLTGTGGLRFNLWSSESA-AWVGQFDESKLPLEQFINWKKYK 195
 DB 161 -VLKHT--ATTNIPKTPGQIMNMLNGTGVSDWLGSYNGVN-PLYAEYDWMV--KYT 210

RESULT 11

Q07856 PRELIMINARY; PRT; 237 AA.

AC Q07856: 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE Beta-(1,3-1,4)-glucanase precursor.
 GN LICHENASE.
 OS Streptococcus bovis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID=1315;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JBI;
 RA Eklund M., Flint H.J.;
 RT "Isolation and overexpression of a gene encoding an extracellular
 beta-(1,3-1,4)-glucanase from Streptococcus bovis JBI."
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RL EMBL: Z92911; CAB07443.1;
 DR HSSP: P23904; IAKJ.

Query Match	12.6%	Score 230;	DB 2;	Length 237;
Best Local Similarity	33.1%;	Pred. No. 5.2e-08;		
Matches	54;	Conservative	26;	Mismatches 69; Indels 14; Gaps 6
OY	30	FSGAELTYLEVOYGFARAKKMAASGTVSSMFLYONGSELTADGRPVYEVILEVIGKRP	89	
DB	81	YTGGEWRKSRKFRFGGLFQVKNKPKIKNPVSSFFTYTGPS--DETGMDEIDIEFLGKDT	137	
OY	90	GSPFSNITTGKAGAKQKTESEKHNHVAVRADQAHTYGLTEPRYVWMTYVDGCEVRTEGGQ	149	
DB	138	TKVGFENYTT----SGGNHEHYLYLNLGFDASQGHHTYGFQMDHITMYVDGRAVYTA-----	190	
OY	150	VSNTLGTGQ-LRFNLV--SSSAWVGVGFDESKRLPLFQFINVY	189	
DB	191	YNNIPSTPGKIMNMAMPCTHEVDSWLGAYN-GRTELYAYYDWI	232	
RESULT 12				
ID	0930U11	PRELIMINARY;	PRT;	205 AA.
AC	0930U11;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	Endo 1-3,1-4-beta-glucanase (fragment).			
OS	Uncultured bacterium.			
OC	Bacteria; environmental samples.			
OX	NCBI_TaxID=77133;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Rodriguez V., Mellado R.P.;			
RT	"Novel lichenases from soil."			
RL	Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: AF254960: AAK50613.1: -			
DR	InterPro: IPR000757; Glyco_hydro.16.			
DR	Pfam: PF00722; Glyco_hydro.16; 1.			
DR	PROSITE: PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.			
FT	NON_TER	1		
FT	NON_TER	1		
FT	NON_TER	1		
SO	SEQUENCE	205 AA;	22890 MW;	E7CADAFE02A2F51 CR664;
Query Match	12.2%;	Score 223;	DB 2;	Length 205;
Best Local Similarity	36.1%;	Pred. No. 1.3e-07;		
Matches	56;	Conservative	21;	Mismatches 68; Indels 10; Gaps 5
OY	23	TNVSAKDPSGAEIYLYLEVOYGFARAKKMAASGTVSSMFLYONGSELTADGRPVYEVII	82	
DB	52	TSPSNKPKDCGEMRSVQTYGYGLYEVVRMKPAKNTGIYVSFFTYTGPE--GTPWDEIDI	108	
OY	83	EVLGNPSPFSQSNITTKGAKAQKTESEKHNHVAVRADQAHTYGLTEPRYVWMTYVDGCEV	142	
DB	109	ESLGGDTTKVQFVNYTNGAG--NHEKTLADLGFDAANAHTMAFTDQPSGIVMYVDGQ-L	164	
OY	143	RKTEGGVSNLGTQGLRFNLWSSESA-AWVGQF	176	
DB	165	KHTATQYPAAPGK--IMMNLWNGGVDDMLGSYN	197	
RESULT 13				
ID	097FD3	PRELIMINARY;	PRT;	246 AA.
AC	097FD3;			
DT	01-OCT-2001 (TREMBLrel. 18, Created)			
DT	01-OCT-2001 (TREMBLrel. 18, Last sequence update)			

Query Match	12.1%	Score 221	DB 16	Length 246
Best Local Similarity	33.7%	Pred. No. 2.2e-07		
Matches 58	Conservative 19	Mismatches 73	Indels 22	Gaps 8
Db	30	FSGAEYLTLAEVQYGFKEARMKMAASGVSSMFLYQNSSEIADGPRWVEDIEVLGKNP	89	
Qy	90	GSFQSNITITKXAGAKQTSSEK---HAVSEPADQAEFTYGLQEMTPYVRYTVDGOEYRKTE	146	
Db	91	YAGGFTSRNNKRYGVLGRVSMKPAKHIGVDSSFFSYTGCS---DNNPWEDIEFLGKOT	147	
Qy	148	TEVQCNYYTNTNG-----KHLELYKLKGIDASKGFHTYGYIMEQNTIALVDGKEYYRA-	200	
Db	147	GGQVSNLTGTGCG-LRFNLMSSESA-AWVGQFDESKLPLEFQTLNWKVRYTTP	196	
Qy	201	--TSNIPTHPGKVMNMLMPCIGDVSWLGAYD-GVTPVYKAYYMAA---MYNP	245	

RESULT 14

Q808N5 PRELIMINARY: PRT: 289 AA.

AC Q808N5; ID 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Endo-1,3-1,4-beta-glycanase.

GN EXOK OR ATU04055 OR AGR_L1600.

OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Rhizobium.

OX NCBI_TaxID=176299;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21608550; PubMed=11743193;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P., Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L., Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr., Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C., Kutyavint T., Levy R., Li M.-J., McClelland E., Palmeri A., Raymond C., Rouse G., Sengenbimmach C., Wu Z., Romero P., Gordon D., Zhang S., Yoo H., Tao Y., Biddle F., Jung M., Krespan W., Perry M., Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M., Chunley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V., Nester E.W.;

RA "The genome of the natural genetic engineer Agrobacterium tumefaciens C58."

RT Science 294:2317-2323(2001).

RL [2]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 10:52:19 ; Search time 15 Seconds
(without alignments)
684.573 Million cell updates/sec

Title: US-09-654-652A-3
Perfect score: 1824
Sequence: 1 MNIKTAVKSAALAVAAAAA.....AKGAKVNPNGHKRRVNFPH 349

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2-6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2-6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2-6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2-6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2-6/ptodata/1/1aa/PTOS.COMB.pep:*
6: /cgn2-6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1210	66.3	228	3	US-09-286-690-12 Sequence 12, Appl
2	248	13.6	239	1	US-08-103-998-2 Sequence 2, Appl
3	248	13.6	242	3	US-09-286-690-8 Sequence 8, Appl
4	233.5	12.8	243	3	US-09-286-690-10 Sequence 10, Appl
5	231.5	12.7	279	3	US-09-286-690-9 Sequence 9, Appl
6	227.5	12.5	238	3	US-09-286-690-7 Sequence 7, Appl
7	227.5	12.3	242	3	US-09-286-690-11 Sequence 11, Appl
8	222.5	12.2	237	1	US-08-103-998-4 Sequence 4, Appl
9	190.5	10.4	245	3	US-09-286-690-2 Sequence 2, Appl
10	159.5	8.7	306	2	US-08-824-707-2 Sequence 2, Appl
11	145	7.9	263	4	US-09-159-106-2 Sequence 2, Appl
12	144	7.9	435	4	US-09-159-106-11 Sequence 11, Appl
13	141.5	7.8	303	4	US-09-159-106-13 Sequence 13, Appl
14	133.5	7.3	276	2	US-08-712-072C-3 Sequence 4, Appl
15	119	6.5	321	2	US-08-712-072C-3 Sequence 3, Appl
16	117.5	6.4	289	2	US-08-580-545B-4 Sequence 4, Appl
17	117.5	6.4	289	4	US-09-262-653A-4 Sequence 4, Appl
18	117.5	6.4	289	4	US-08-867-484A-2 Sequence 2, Appl
19	117	6.4	620	4	US-09-269-731-6 Sequence 6, Appl
20	116.5	6.4	634	1	US-08-392-828C-2 Sequence 2, Appl
21	116.5	6.4	654	1	US-09-330-945-2 Sequence 2, Appl
22	115	6.3	262	1	US-08-392-828C-37 Sequence 37, Appl
23	115	6.3	262	3	US-09-330-945-37 Sequence 37, Appl
24	114.5	6.3	545	4	US-09-269-731-8 Sequence 8, Appl
25	113	6.2	285	2	US-08-712-072C-5 Sequence 5, Appl
26	110.5	6.1	295	1	US-08-295-657-3 Sequence 3, Appl
27	110	6.0	287	4	US-08-640-737-8 Sequence 8, Appl

28	109.5	6.0	289	4	US-08-640-737-2	Sequence 2, Appl
29	109	6.0	382	4	US-08-213-419B-19	Sequence 19, Appl
30	108.5	5.9	367	4	US-08-213-419B-15	Sequence 15, Appl
31	108.5	5.9	989	4	US-08-213-419B-2	Sequence 2, Appl
32	108.5	5.9	989	4	US-08-213-419B-4	Sequence 4, Appl
33	107.5	5.9	583	4	US-09-620-412C-353	Sequence 353, App
34	107.5	5.9	666	4	US-08-961-083-2	Sequence 2, Appl
35	106.5	5.8	682	3	US-08-481-435-6	Sequence 6, Appl
36	105	5.8	422	2	US-08-712-072C-2	Sequence 2, Appl
37	104	5.7	348	4	US-09-216-295-16	Sequence 16, Appl
38	103.5	5.7	395	2	US-08-404-531B-3	Sequence 3, Appl
39	103.5	5.7	395	3	US-08-476-900A-3	Sequence 3, Appl
40	103.5	5.7	395	3	US-08-488-546A-3	Sequence 3, Appl
41	103.5	5.7	1580	4	US-08-726-320-1	Sequence 1, Appl
42	103.5	5.7	1580	4	US-09-208-716-1	Sequence 1, Appl
43	103.5	5.7	1581	4	US-08-726-320-3	Sequence 3, Appl
44	103.5	5.7	1581	4	US-09-208-716-3	Sequence 3, Appl
45	98.5	5.4	1581	2	US-08-404-531B-6	Sequence 6, Appl

ALIGNMENTS

```
RESULT 1
US-09-286-690-12
: Sequence 12, Application US/09286690
: Patent No. 6103511
: GENERAL INFORMATION:
: APPLICANT: Li, Xin-Liang
: APPLICANT: Ljungdahl, Lars G.
: APPLICANT: Chen, Huizhong
: TITLE OF INVENTION: Lichenase and Coding Sequences
: FILE REFERENCE: 55-56
: CURRENT APPLICATION NUMBER: US/09/286,690
: CURRENT FILING DATE: 1999-04-05
: EARLIER APPLICATION NUMBER: US 60/027,882
: EARLIER FILING DATE: 1999-10-04
: EARLIER APPLICATION NUMBER: PCT/US97/17811
: EARLIER FILING DATE: 1997-10-03
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: Patentn Ver. 2.0
: SEQ ID NO 12
: LENGTH: 228
: TYPE: PRT
: ORGANISM: Fibrobacter succinogenes
: US-09-286-690-12

Query Match      66.3%; Score 1210; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.3e-102;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MNIKTAVKSAALAVAAAAALTTNVSAKPSCAELTYLLEVOYGFEMKMAASGTVS 60
DB      1 MNIKTAVKSAALAVAAAAALTTNVSAKPSCAELTYLLEVOYGFEMKMAASGTVS 60
QY      61 SNEVLONGSEIDGRPMVEVDIEVLCKNPGSPQSNITTKAGAKOKTSEKHHAVSPAADA 120
DB      61 SNEVLONGSEIDGRPMVEVDIEVLCKNPGSPQSNITTKAGAKOKTSEKHHAVSPAADA 120
QY      121 FHTYGLEWTPNVRMTVDQEVKRTGEGVSNLTGQGLRFLNMSSESAAMVGQFDESKL 180
DB      121 FHTYGLEWTPNVRMTVDQEVKRTGEGVSNLTGQGLRFLNMSSESAAMVGQFDESKL 180
QY      181 PLFOPLNWKVKYKYPGQEGSGDFLDMWTDFDGRMKGDWTF 228
DB      181 PLFOPLNWKVKYKYPGQEGSGDFLDMWTDFDGRMKGDWTF 228

RESULT 2
US-08-103-998-2
: Sequence 2, Application US/08103998
: Patent No. 5470725
: GENERAL INFORMATION:
```

```

; APPLICANT: Borris, Rainer
; APPLICANT: Hofemeister, Jürgen
; APPLICANT: Thomsen, Karl Kristian
; APPLICANT: Olsen, Ole
; APPLICANT: Von Weltstein, Dietrich
; TITLE OF INVENTION: A thermostable (1,3-1,4)-beta-Glucanase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 1800 Diagonal Road, P.O. Box 299
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22113-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/103,998
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/773,652
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30307/123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-103-998-2

Query Match      13.6%; Score 248; DB 1; Length 239;
Best Local Similarity 35.4%; Pred. No. 6.8e-15;
Matches 62; Conservative 27; Mismatches 70; Indels 16; Gaps 8;

QY 23 TNVSAKDSGAEIYLYEEVQYGFARMKMAASGTVSMFLYONGSEIADGRPAVEYDI 82
DB 77 TSPSYNKFDCGNNRSVQYTGYLEVRMKPAKNTGIVSSFTYTGPT--GTPWDEIDI 133
QY 83 EVLGKNNPGSFQSNITITGKGAQKTSEKHHAASPADDAFHHTYGLEMTPNYVRWTVDGOEV 142
DB 134 EFLGKDTTKVOGFNYTNGVGH---EKYISLGFDAKSGFHTYAFDMQNSIKWYVDG-VL 189
QY 143 RKTGGQVSNLTGTGCG-LRFNLMSSEA-AWVGQFDESKLPLEFQINWYKYVK 195
DB 190 KHT--AVANIPSTPGKIMMNLNMGTVDMWLGSYNGAN-PLXAEYDWV---KYT 237

RESULT 3
US-09-286-690-8
; Sequence 8, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04

```

```

; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Bacillus subtilis
; US-09-286-690-8

Query Match      13.6%; Score 248; DB 3; Length 242;
Best Local Similarity 36.0%; Pred. No. 6.9e-15;
Matches 63; Conservative 25; Mismatches 75; Indels 12; Gaps 7;

QY 20 ALTNVSAKDSGAEIYLYEEVQYGFARMKMAASGTVSMFLYONGSEIADGRPAVEYDI 79
DB 78 ALTPSAVYK-FDCGNNRSVQYTGYLEVRMKPAKNTGIVSSFTYTGPT--DGTWDE 133
QY 80 VDEVLGKNNPGSFQSNITITGKGAQKTSEKHHAASPADDAFHHTYGLEMTPNYVRWTVDGO 139
DB 134 IDIEFLGKDTTKVOGFNYTNGV---NHEKIYDLGFDANAHHTYAFDMQNSIKWYVDG 190
QY 140 QEVKRTGGQVSNLTGTGCLRFNLMSSEA-AWVGQFDESKLPLEFQINWYKYVK 193
DB 191 Q-LKHTATNQTPTTPEK--IMMNLNMGTVDEMLGSYNGV-PLXAEYDWVRYTK 241

RESULT 4
US-09-286-690-10
; Sequence 10, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
; US-09-286-690-10

Query Match      12.8%; Score 233.5; DB 3; Length 243;
Best Local Similarity 33.7%; Pred. No. 1.5e-13;
Matches 58; Conservative 26; Mismatches 77; Indels 11; Gaps 6;

QY 23 TNVSAKDSGAEIYLYEEVQYGFARMKMAASGTVSMFLYONGSEIADGRPAVEYDI 82
DB 77 TSPSYNKFDCGNNRSVQYTGYLEVRMKPAKNTGIVSSFTYTGPT--DGTWDEIDI 137
QY 83 EVLGKNNPGSFQSNITITGKGAQKTSEKHHAASPADDAFHHTYGLEMTPNYVRWTVDGOEV 142
DB 138 EFLGKDTTKVOGFNYTNGV---NHEKIVNLGFDPAANSYHTYARDWQNSIKWYVDG-L 193
QY 143 RKTGGQVSNLTGTGCLRFNLMSSEA-AWVGQFDESKLPLEFQINWYKYVK 193
DB 194 KHTATQIIPQTPGK--IMMNLNMGAGVDEMLGSYN-GVTPLSRSLHWVRYTK 242

RESULT 5
US-09-286-690-9
; Sequence 9, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:

```

APPLICANT: Li, Xin-Liang
 APPLICANT: Ljungdahl, Lars G.
 APPLICANT: Chen, Huizhong
 TITLE OF INVENTION: Lichenase and Coding Sequences
 FILE REFERENCE: 55-96
 CURRENT APPLICATION NUMBER: US/09/286,690
 CURRENT FILING DATE: 1999-04-05
 EARLIER APPLICATION NUMBER: US 60/027,882
 EARLIER FILING DATE: 1999-10-04
 EARLIER APPLICATION NUMBER: PCT/US97/17811
 EARLIER FILING DATE: 1997-10-03
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 9
 LENGTH: 279
 TYPE: prt
 ORGANISM: Clostridium thermocellum
 IS-09-286-690-9

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/103,998
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/773,652
;; FILING DATE: 16-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bent, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 30307/123
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-9300
;; TELEFAX: (703) 683-4109
;; TELEX: 899149
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 237 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-103-998-4

Query Match 12.2%; Score 222.5; DB 1; Length 237;
Best Local Similarity 33.1%; Pred. No. 1.4e-12;
Matches 57; Conservative 25; Mismatches 79; Indels 11; Gaps 6;

QY 23 TVNSAKDFSGAELTYLEEVQYGFARMKMAASGVSSMFLYONGSEIADGRPVYVDI 82
DB 75 TSSAANKDCACAEYRSTNYIGLYEVSMPAKNTGIVSSFFVTGP--AHGTQDEIDI 131
QY 83 EVLGNKPGSFQSNITITGKAGAKTSEKHNAVSPADQAFHTYGTGLETPNYVWTVGQEV 142
DB 132 EELGADTTKRYQNFNYTNGAG--NHEKPADLGFDANAHYTAFFDMQPSISINYVDGQ-L 187
QY 143 RKEGGOVSNLTGTGLRPNLWSSSA-AWVGQFDESKLPLEFQINWYKVKYK 193
DB 188 KHTATTOIPAAPGK--IMMNLNMGTVGVDMLGSLYNGVN-PIVAHYDMMYRK 236

RESULT 9
US-09-286-690-2
;; Sequence 2, Application US/09286690
;; Patent No. 6103511
;; GENERAL INFORMATION:
;; APPLICANT: Li, Xin-Liang
;; APPLICANT: Ljungdahl, Lars G.
;; APPLICANT: Chen, Huizhong
;; TITLE OF INVENTION: Lichenase and Coding Sequences
;; FILE REFERENCE: 55-96
;; CURRENT APPLICATION NUMBER: US/09/286,690
;; EARLIER FILING DATE: 1999-04-05
;; EARLIER APPLICATION NUMBER: US 60/027,882
;; EARLIER FILING DATE: 1999-10-04
;; EARLIER APPLICATION NUMBER: PCT/US97/17811
;; EARLIER FILING DATE: 1997-10-03
;; NUMBER OF SEQ ID NOS: 12
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 2
;; LENGTH: 245
;; TYPE: PRT
;; ORGANISM: Oriblomyces sp. PC-2
US-09-286-690-2

Query Match 10.4%; Score 190.5; DB 3; Length 245;
Best Local Similarity 31.6%; Pred. No. 1.2e-09;
Matches 54; Conservative 17; Mismatches 87; Indels 13; Gaps 6;

QY 21 LTTNWSADFSQAELELYLEEVQYGFARMKMAASGVSSMFLYONGSEIADGRPVYEV 80
DB 79 LTIDRDSGYTCGEFRTNRYGYGMFOVNMKPIKPGVSSFFVTGSPS--DGTKMDEI 135
QY 81 DIEVLGNKPGSFQSNITITGKAGAKTSEKHNAVSPADQAFHTYGTGLETPNYVWTVGQ 140

DB 136 DIEFLGYDTTKYQNFNYTNGQGH--EHIHILGFDASGCFHTYFFMARNSITMYVGT 192
QY 141 EVRKEGGOVSNLTGTGQ-LRPNLWSSSA-AWVGQFDESKLPLEFQINWYK 189
DB 193 AVYTA---YDNIPTPKIKIMNANNGICVDMLRPFN-GRTNISAYYDWV 238

RESULT 10
US-08-824-707-2
;; Sequence 2, Application US/08824707
;; Patent No. 5919688
;; GENERAL INFORMATION:
;; APPLICANT: Ferrer, Pau
;; APPLICANT: Diers, Ivan
;; APPLICANT: Hedegard, Lisbeth
;; APPLICANT: Haikier, Torben
;; APPLICANT: Asenjo, Juan
;; APPLICANT: Savva, Demetris
;; TITLE OF INVENTION: No. 5919688e1 enzyme with beta-1,3-glucanase activity
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: No. 5919688o No. 5919688disk of No. 5919688th America, Inc.
;; STREET: 405 Lexington Avenue, Suite 6400
;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10174-6401
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/824,707
;; FILING DATE: 14-April-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gregg, Valeria A.
;; REGISTRATION NUMBER: 35,127
;; REFERENCE/DOCKET NUMBER: 4290, 204-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-867-0123
;; TELEFAX: 212-878-9655
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 306 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-824-707-2

Query Match 8.7%; Score 159.5; DB 2; Length 306;
Best Local Similarity 27.6%; Pred. No. 1.1e-06;
Matches 59; Conservative 29; Mismatches 89; Indels 37; Gaps 11;

QY 2 NIKKTAVKSALAVAAAAALTTNWSAKDFSGAELTYLEEV--OYGFARMKMAASGV 59
DB 103 NITTSRVNSALDGGCNLYITLALDSDGSYTSARLTGQNVNPOFGRIEARLIQIPROGIIW 162
QY 60 SSMFLYONGSEIADGRPV--EVDI-EVLGNKPN---GSFQSNITITGKAGAKTSEKH 111
DB 163 SAFMWV--GANLPLD-TMPPTSGEIDIMENVGNARHEVIGYHAGCYSDNCKIMGTYQHPO 219
QY 112 AVSPADQAFHTYGTGLETPNYVWTVGQEVKRTGGOVSNLTGTGLRPNLWSSSAW 171
DB 220 GMSFADD--FHTFGIDWTPCEITLWLVGQEVHRYTADVG-----ANOW 261
QY 172 VGQFDESKLPLEFQINWYKVKYTPGQEGSSDF 205
DB 262 V--FDQ--PFLLIN-VAIGGQWPGNDATTPF 289

Db 244 WEEDNIKVVYDGKFFYKVTNOQ-----WYSTAAPNNPNAPFDEPEYLIWNL 289
QY 170 AWGQFD-----ESKLPLOFIMWVKYKY 194
| | | | |
Db 290 AVGSNFDGRTPNASDIPATMVDYVRVYKF 320

Search completed: July 11, 2003, 10:54:36
Job time : 17 secs

OY 195 TPGGEGGSDP 205
11 : :
DB 267 WPGDPMSTAF 277

RESULT 2

US-09-734-569-152
Sequence 152, Application US/09734569
Patent No. US20020064816A1

GENERAL INFORMATION:

APPLICANT: Lerchl, Jens
APPLICANT: Renz, Andreas
APPLICANT: Ehrhardt, Thomas
APPLICANT: Reinold, Andreas
APPLICANT: Clippus, Petra
APPLICANT: Bischoff, Friedrich
APPLICANT: Frank, Markus
APPLICANT: Freund, Annette
APPLICANT: Duwenig, Elke
APPLICANT: Schmidt, Ralf-Michael
APPLICANT: Reski, Ralf
TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved in the synthesis of carbohydrates
FILE REFERENCE: BASE-NAME-1332-99-US
CURRENT APPLICATION NUMBER: US/09/734,569
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 60/171,101
NUMBER OF SEQ ID NOS: 181
SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1
SEQ ID NO 152
LENGTH: 287
TYPE: PRT
ORGANISM: Physcomitrella patens
US-09-734-569-152

Query Match 6.6%; Score 120; DB 10; Length 287;
Best Local Similarity 23.0%; Pred. No. 0.031;
Matches 49; Conservative 30; Mismatches 70; Indels 64; Gaps 11;

OY 43 YGKFEARKMAA--ASGVSSMFLYQNGSEIDGRPWVEVDIEVGKNGS---FQSNII 97
DB 69 YVDISATYIKMPEDSAGVYTFYFSSQGDQ-----HYELDEFLGNTSGQPFLLHTNVF 122
OY 98 TGKGAQKQTSKHNHVAAPADQAFHTYGLTPTNYVMTVDQGEVKTGEGGVSNLTGT- 156
DB 123 VDGVGRE-QQMYLGFDSAD--FHYFRMSKDMVYFYVDKPYR-----VFKNLEGT 174
OY 157 -----QGLRFNLMSSESAAVGFDESKLPLFOFINWVKYKXYPGQEGGSDFT 206
DB 175 PGTXYLNOQAMGVYISIDGSSMATQG---GRVP-----INW-----ASAPPT 214
OY 207 LDMTNDFTFDGSRKMGKDMFTFDGRRVLDLTKN 239
DB 215 ATYQ-----DFALNGCVDPNDPN 233

RESULT 3

US-10-156-761-9339
Sequence 9339, Application US/10156761
Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 9339
LENGTH: 629
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-9339

Query Match 6.4%; Score 117.5; DB 9; Length 629;
Best Local Similarity 23.4%; Pred. No. 0.14;
Matches 47; Conservative 34; Mismatches 69; Indels 51; Gaps 9;

OY 29 DFGS--AEIYLYEVQYGFARMKMAASGVSSMFLYQNGSEIAD--GRPWVEVDIEVL 85
DB 438 DFTSGRVDTNTRFEDTYGRVSARMLPVODGCFWPAFWLL--GSNVDDPSVSW----- 487
OY 86 GKNPGSFOSNITGKAGAKTSEKHNAVSPAD-----QAFTYGLTPT 130
DB 488 ---PASGETDWMENIGYSDMTSSALHGPYSADGICARQYTPGGGTADHMTHTYVWTP 544
OY 131 NYVRMTVDQGEVKTGEGGVSNLTG-----TQGLRFNL-----WSSESAAVGC-- 173
DB 545 TMRFTVDDRLVQETFRNKLESTRGQWYDHNQYVILNLALGCAVPAGMNQVTSFYWGLP 604
OY 174 QPDESKLP---LFOFINWVKV 191
DB 605 QTSVDKVAAGVQAEVDWVRV 625

RESULT 4

US-09-988-200-6
Sequence 6, Application US/09988200
Patent No. US2002009453A1

GENERAL INFORMATION:

APPLICANT: BARBEYRON, Tristan
POTIN, Philippe
RICHARD, Christophe
HENRISSAT, Bernard
YVIN, Jean-Claude
KLOAREG, Bernard
TITLE OF INVENTION: Glycolyse hydrolase genes and their use for producing enzymes for the biodegradation of Carriageans

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: DENNISON, MESEROLE, SCHEINER & SCHULTZ

STREET: 612 Crystal Square 4, 1745 Jefferson Davis

Highway

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/988,200

FILING DATE: 19-NO. US2002009453A1-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/269,731

FILING DATE: <unknown>

APPLICATION NUMBER: FR 96 12204

FILING DATE: 07-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: IRA SCHULTZ

REGISTRATION NUMBER: <unknown>

```

QY      174 QFD-ESKLP LQF INMVKVYKYTPGQGBG-----SDFTLMDWTDNFDYED 217
      11 :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

Db 142 WLYSDFDYSVANGETVYSEIDVV

142 WLSDFDYSVANGETVYSEIDVV

Db 202 KMYPQEQLLKMKRAMDPSKD--FYHIGCEVQNQNEIILWYDVEARKKPKYWHRRMNVLTLS 259

Qy 157 QGLR-----FNIMSSESAANYGQFDE-----SKLPLEQFINWVKYKYYKTTPQGGEGSGPFTL 207

Db 260 LGRKRPPVKFEDKNKNAIPIETDAKAREKLSIDIPSMVDYVRAWEKSAGN----- 310

Qy 208 DMTDNEDTFD-----GSRMGKGDMTEPDGNVDTLDITNIIYSRDGMILLALTTRKQGESFN 260

Db 311 --TTNPPTSEVGLTKTKGSKLVIDHMDASTGISAIVSNN-----TKTGO--YA 354

Qy 261 GQVPRDDEPAQSSSSAPASSSSVASSSSVAPASSSASFVPPSS-----SGATN 309

Db 355 GSV-----NNASIQIYTLKAAISYKVASGAKSSPQTSYVLGISAASN 398

Db 27 IGAMATAAGDILKPLTHVIGPTAFLG-----LGVDNNGNGARVGVGAPAPASLGI-ST 80
QY 68 GSEL--ADGRP---WEVEVDIEVLCKNGSPQSNITTKAGAKOQTSKHHAVSPADQAFH 122
Db 81 GDVITFADGAPINSATFAMADALNGHHPGDVISTVMQKSGTFRGNVTLLEGPAE--FC 138
QY 123 TYGLEWPNVVRMTVDGQEVKTE-----GG-----QVSNLTGQGLFNL 163
Db 139 RYSHMRP-----LDTQVSESPESPTSPDDVLAKGGGITYETKSLTITGTLIDVSN 192
QY 164 WSESAAWVGQFDESKLPFQFINWVYKRTPGQGGSDFTLDMTDNFTPDGSRWGK 223
Db 193 -ATDSGA--GVTFKRNLSCTN-TNSLQFLKNSAGHGGAIVYQTMSTVTSTSTTPP 248
QY 224 --GDMTDRNRVDLTKR-----NYSRDKMLIALTKRQGSFNG-----QVPRD 266
Db 249 LVGEVIFSEN---TAKHGCGICTNLSLNKTVTLTNSAKESGGAFTDLASIPPT 304
QY 267 DEBAPQ--SSAPASSSSVPAS-----SSSVPASSSAFVPPSSSATNAIGHMRTT 317
Db 305 DTBESSIPSSSSPASTPEVVASAKINRFASTAPAPPS---LTAESDQTDQETSDTN 361
QY 318 PAVAKHRLVNA-----KGAKVNPNGHKRRVYN 346
Db 362 SDIVSTENILNVAINQNTSAKKGALYKKAKLSRIN 399

RESULT 10
US-09-765-272-2
; Sequence 2, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-765-272-2
Query Match 5.9%; Score 107.5; DB 10; Length 666;

Best Local Similarity 23.5%; Pred. No. 1.1;
Matches 76; Conservative 50; Mismatches 104; Indels 93; Gaps 20;
QY 57 GTVSSMFLYNGSEFIADGRPWVEVDIEVLCKN-----PGSFQSNITTKAGA 103
Db 364 GNITLQYALQCSRNV---PAVET-LNKVGLNRAKTEPLNGLIDIDPSIHNSMIS--SNT 416
QY 104 OKTSEKHAVS---PADQAFHITYGLEMTNRYRWTV--DGOEVKTEGGOVSNLTGTOG 158
Db 417 TESDKRYGASSEKMAAAYAFANGGYKKPMYIHKRVFSQSE-----KEFSNV-GTRA 469
QY 159 LRFNLMSESAWVGQFDESKLPF-----QFINWVYKRTPGQGE--GSGDFTLDWTD 211
Db 470 MK-----ETTAIV-MTDMKTYTLTYGTGNAYLAWL-----PQAKTQTSNVTDEIE 516
QY 212 NF-----DTFGD--SRWKGDMTFDGNRY-----DLIDKNYSRDKMLILA--- 250
Db 517 NHKTSQFVAHDELFAGYTRKYSMAVWTGYSNRLTPLVGNQLVAARVYSMMTYLSEGS 576
QY 251 -----LTKRQGSF--NGQVPRDDEBAPQ-----SSSAPASSSSVPASSSVPASS 295
Db 577 NPEDWNIPEGILYRNGEVEFKNGARSTWNSPAPQDPSTESSSSSDSSTSOSSSTTPSTN 636
QY 296 SSAFVPPSSSATNAIGHMRTTP 318
Db 637 NSTTTNPNMNRQGS-----NTTP 654

RESULT 11
US-09-815-242-13423
; Sequence 13423, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13423
Query Match 5.8%; Score 106.5; DB 10; Length 719;
Best Local Similarity 23.5%; Pred. No. 1.5;
Matches 76; Conservative 49; Mismatches 105; Indels 93; Gaps 20;
QY 57 GTVSSMFLYNGSEFIADGRPWVEVDIEVLCKN-----PGSFQSNITTKAGA 103

```
Db 417 GNITLQVLAQOSRNV-----PAVET-LNKVGLNBAKTFNLGLGIDVPSIHYSNALS--SNT 469
Oy 104 OKTSKHHAVS---PADQAHTHTGLEWTPNVVRATV--DOQEVAKTGGQVSNLTGOC 158
Db 470 TESDKKYGASSSEKMAAAAFANGGTYYKPMYIHKVYVESDSE-----KFEFNV-GTRA 522
Oy 159 LRFNLMSSESAAMVGOFPDESKLPLF-----QFINMVVYKYTPQOGE-GGSDFTLDWTD 211
Db 523 MK-----ETTAIV-MTDMKMTVLSTYGRNAIYLAWL-----PQAGKTGTSNTDEIE 569
Oy 212 NF-----DTEFDG--SRMGKGDWTFDGNRV-----DLTDBKNYSRDCMLILA--- 250
Db 570 NHIKTSQFVADDELFACTYRKYSMAVMTGYSNRLTPLVNGLTVAAKVYRSMMTYLSGGS 629
Oy 251 -----LTKRGOESF-NGQVPRDEPAPO-----SSSAPASSSSVPASSSSVPASS 295
Db 630 NPEDMNIEGLYRKNGEYEVFKNGANSTWSSPAQOPPSTESSSSSDSTSCSSSTTPTSTN 689
Oy 296 SSAFVPPSSSATNAINHGMRTTP 318
Db 690 NSTTTPNNTNTOQS-----NTTP 707
```

RESULT 12

US-10-156-761-11953

Sequence 11953, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 11953

LENGTH: 698

TYPE: PRT

ORGANISM: Streptomyces avermitilis

US-10-156-761-11953

Query Match 5.8%; Score 106; DB 9; Length 698;

Best Local Similarity 20.1%; Pred. No. 1.5; Indels 139; Gaps 20;

Matches 79; Conservative 50; Mismatches 126; Indels 139; Gaps 20;

Db 24 NVSAKDFSGAB-LYTLLEVQYKFEARKMAAASGVSMFLYONGSEIADGRPWEVDI 82

Db 180 NIGSLDLSGADPPVTLKO-----DINRMW--YDA 206

Oy 83 EVLGKN---PGSF-----GSNITTC--KACAQKTSKHHAVSPAADOAFHTYGLEWTPN 131

Db 207 PILSSSAGAPGTLVAGAPGOSPVLGVYDVSSGTATLTAHAFFDEGNTGGGLSLAVTP- 265

Oy 132 YVRMTVGOQEVKRTGEG-----QVSNLTGTOGLRFLMW-----SSE9AAMVGOFPDES 178

Db 266 -----DGDVVTASGAPYIOAVYKISDLSADGRYQTVTPNAVDIAPDGTVAAGTF--- 316

Oy 179 KLPLFOFINMV--KVYKXTPQOGEKGSDFTLDMT---DNFDPDGRKMGKDMWTFDGNR- 232

Db 317 -----SWYDPVHVHFKPGVSTPRVQYDLRPNGTSSGADTLVLAGSGLA---NSPDGSR 365

Oy 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

Db 233 --VDLTKNITS-----RDGMLILA---LTKRGOESFNGQVPRDEPAPOSSSSA 277

```
Db 366 FAVSYNDKDYSLRFDPPFKATSLTMVAVPKATRGKALSVSGLT-----TSAEAL 417
Oy 278 PA-----SSSVPASSSSVPASSSSAF--VPP-----SSSS 306
Db 418 PAGATVSVTRTDMESPCKPVGATVAAADGTYTGDPAGCKVYATATYAGASHGAT 477
Oy 307 ATNAINHGMRTTPAAVAKHEHRLVNAKGAQVNPNGH 340
Db 478 ATGSVEVSRSTPTLTLLNNDGKVYAYGSDVKRTAH 511
```

RESULT 13

US-10-156-761-9706

Sequence 9706, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 9706

LENGTH: 311

TYPE: PRT

ORGANISM: Streptomyces avermitilis

US-10-156-761-9706

Query Match 5.6%; Score 102.5; DB 9; Length 311;

Best Local Similarity 22.3%; Pred. No. 1; Indels 101; Gaps 17;

Matches 84; Conservative 44; Mismatches 148; Indels 101; Gaps 17;

Db 1 MNIRKTAVKSALVAAAAALTTNVSAKDFSGAELTYLLEVQYKFE--ARMKMAASGT 58

Db 1 MTAHRTAAAVAAVAAPELLLTMAAGPAQAAGAPTDPVSRVACSPGCGRARTAACTAA 60

Oy 59 VSSMFLVONGSEIADGRPWEVD-IEVLGNKPGSFQ-----SNITGKAQAKTSEKHHAV 113

Db 61 IA-----ANGSPFTAMDNLRVAGVGRDRDOLIPDGKLCGGILPYRK----- 101

Oy 114 SPADAQAFHTYGLEWTPNRYVMTVDQGEVRRKTEGGQVSNL-----TGTQGLRFLNLS 166

Db 102 -----GLDLTNS-----DWPSTRLTTPASLTMTKSSITPTGT----FKLYLT 140

Oy 167 ESAAMVGOFPDESK-----LPLFOFINMVKVYKXTPQOGEKGSDFTLDMTDNFDPDGR 220

Db 141 KQG-----YDPTKPLTMSDLPAPQFAQ-LKDPALAGAYRLGAKLPADPRGRHVLF--TI 192

Oy 221 WGRKDWTFDGNRVULTKRNITSRDCMLILALTRKGOESFNGQVPRDEPAPOSSSSAPAS 280

Db 193 W-----QNTSTDTLTYYSCSD--VVFKAKAAMAAGAGSGTCAARKPAGKATPSSPAA 241

Oy 281 SSSVPASSSSVPASSSAFVPPSSSATNAINHGMRTTPAAVAKHEHRL----- 326

Db 242 TSS-PSSSPSPSRKETASTPSPSRKAVPG-----TPAATAHRSDSGSLPLLAGAA 294

Oy 327 -LVNAKGAQVNPNGHR 342

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

Db 295 AVVLAAGAVVSLRRRR 311

```

Sequence 10925, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 10925
LENGTH: 904
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-10925

Query Match          5.6%; Score 102; DB 9; Length 904;
Best Local Similarity 21.5%; Pred. No. 4.7;
Matches 59; Conservative 38; Mismatches 126; Indels 52; Gaps 10;

OY 52 MAASGVVSMFLQNG---SEIDGRPWVVDIEVLGKNPQSGNSNIITKAGAQKTSE 108
Db 29 VAGGAGVYTVAVASPSGDADVDVARGKRPVVDLAL-----KAGTSGSRE 74
OY 109 KHAASPAADQAFHTYGLFEMTPNVRMTVDGQEVKTEGGVSNLTGTQGL----- 159
Db 75 ---LATTTTEAFSTVGVSWTC--AARELDGTAGVTRGTGTGEMSGKRNLAPLGVDR 128
OY 160 ---RFINMSSESAAWVGOFDESKLPLQFIWVKYKTYTPQGGEG--SDFTLDWTDNFD 214
Db 129 EPGAKNARGASDPLWVGPSDVGQARVLA-----ADGSARAGLPEGLEVINLVDEGV 178
OY 215 TFDOSRRGKG-GDMTFDGNRVLTGKNKINSRDCMILLATIRKGQSEFNQVPRDDPPAQ 273
Db 179 TTTEAR-NKGLDITSGTGIGDLSNAAPYAKDPQSOTPTGTSPGSATETGSEP--APAGD 234
OY 274 SSSAPASSSSVPASSSSVPASSSSAFVPPSSSSSAT 308
Db 235 SDSASATATATSTATTATTATPAPASTVAPRPSIVSRT 269

RESULT 15
US-10-156-761-9303
Sequence 9303, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 9303
LENGTH: 466

```

[illegible]

PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX
PS Claim 20: SEQ ID NO 55750: 103pp: English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
XX
SQ Sequence 361 AA:

Query Match 3.28; Score 11; DB 22; Length 361;
Best Local Similarity 100.0%; Pred. No. 0.092;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 SSSSAPASSSS 283
Db 135 SSSSAPASSSS 145
|||||
ABP03676 standard; Protein: 88 AA.
ID ABP03676:
XX
AC ABP03676:
XX
DT 25-JUN-2002 (first entry)
XX
DE Human ORFX protein sequence SEQ ID NO:7334.
XX
KW Human: open reading frame; ORFX: gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.
XX
OS Homo sapiens.
XX
PN MO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001MO-US10836.
XX
PR 30-MAY-2000; 2000US-206132P.
PR 29-AUG-2000; 2000US-228716P.
XX
XX (CURA-) CURAGEN CORP.
XX PA Shinketsu RA, Leach MD;
XX PI
XX DR WPI; 2002-106308/14.

DR N-PSDB; ABN19428.
XX
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders -
XX
XX
PS Disclosure; SEQ ID 7334; 1037pp: English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27322 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, Keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
XX
SQ Sequence 88 AA:

Query Match 2.98; Score 10; DB 23; Length 88;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 SSSVPAASSSS 290
Db 60 SSSVPAASSSS 69
|||||
AAG45466 standard; Protein: 144 AA.
ID AAG45466:
XX
AC AAG45466:
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 57085.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
OS
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136332.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138340.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140941.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145228.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155436.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159328.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159568.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160988.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match Best Local Similarity 2.98; Score 10; DB 21; Length 144;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 SALAVAAAA 19
Db 5 SALAVAAAA 14

RESULT 4
AAC45465
ID AAC45465 standard; Protein: 209 AA.

XX AAC45465;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 57084.

XX Protein Identification; signal transduction pathway; metabolic pathway;
KW hydrolisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.

```
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.9%; Score 10; DB 21; Length 209;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 SALAVAAAAA 19
DB 70 SALAVAAAAA 79

RESULT 5
AAC49854
ID AAC49854 standard; Protein; 268 AA.
AC AAC49854;
XX
XX 18-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63112.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX OS
XX Arabidopsis thaliana.
XX PN
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
```

PR 06-MAY-1999: 99US-0132487.
PR 07-MAY-1999: 99US-0132863.
PR 11-MAY-1999: 99US-0134256.
PR 14-MAY-1999: 99US-0134318.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134221.
PR 14-MAY-1999: 99US-0134370.
PR 18-MAY-1999: 99US-0134768.
PR 19-MAY-1999: 99US-0134941.
PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139219.
PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139750.
PR 21-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 06-JUL-1999: 99US-0142055.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.

PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145913.
PR 27-JUL-1999: 99US-0145918.
PR 27-JUL-1999: 99US-0145919.
PR 28-JUL-1999: 99US-0145951.
PR 02-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.
PR 02-AUG-1999: 99US-0146389.
PR 03-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147204.
PR 04-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147192.
PR 05-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147493.
PR 09-AUG-1999: 99US-0147935.
PR 10-AUG-1999: 99US-0148171.
PR 11-AUG-1999: 99US-0148319.
PR 12-AUG-1999: 99US-0148341.
PR 13-AUG-1999: 99US-0148565.
PR 13-AUG-1999: 99US-0148684.
PR 16-AUG-1999: 99US-0149368.
PR 17-AUG-1999: 99US-0149175.
PR 18-AUG-1999: 99US-0149426.
PR 20-AUG-1999: 99US-0149722.
PR 20-AUG-1999: 99US-0149723.
PR 20-AUG-1999: 99US-0149929.
PR 23-AUG-1999: 99US-0149902.
PR 23-AUG-1999: 99US-0149930.
PR 25-AUG-1999: 99US-0150566.
PR 26-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 30-AUG-1999: 99US-0151080.
PR 31-AUG-1999: 99US-0151303.
PR 01-SEP-1999: 99US-0151438.
PR 07-SEP-1999: 99US-0151930.
PR 10-SEP-1999: 99US-0152363.
PR 13-SEP-1999: 99US-0153070.
PR 15-SEP-1999: 99US-0153758.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.

PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161320.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 29-OCT-1999; 990S-0162142.

Query Match 2.98; Score 10; DB 21; Length 268;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SALAVAAAA 19
Db 5 SALAVAAAA 14

RESULT 6
AAG19454
ID AAG19454 standard; Protein; 319 AA.
XX
XX AAG19454;
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 21262.
DE
XX Arabidopsis thaliana protein fragment SEQ ID NO: 21262.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
KM
OS Arabidopsis thaliana.
XX
XX
PN EP1033405-A2.
PD
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.

PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139859.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142877.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144684.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.9%; Score 10; DB 21; Length 319;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 SALAVAAAA 19
Db 5 SALAVAAAA 14

RESULT 7
AAG49849
ID AAG49849 standard; Protein; 322 AA.

XX AAG49849;
AC
XX 18-OCT-2000 (first entry)
DT
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 63105.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

XX
PN EP1033405-A2.

XX
PD 06-SEP-2000.

XX
PF 25-FEB-2000; 2000EP-0301439.

XX
PR 25-FEB-1999; 99US-0121825.
PR 09-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137722.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143622.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147302.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0153363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154772.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0158293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 2.9%; Score 10; DB 21; Length 322;
Best Local Similarity 100.0%; Pred. NO. 0.72;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 SALAVAAAAA 19
|||||||
DB 5 SALAVAAAAA 14

RESULT 8
AAG49853
ID AAG49853 standard; Protein: 333 AA.
XX AAG49853;
XX
XX 18-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63111.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
PN EPI033405-A2.
XX
XX 06-SEP-2000.
PD
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134236.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140693.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144003.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.

PR	10-AUG-1999;	9905-0148171.
PR	11-AUG-1999;	9905-0148319.
PR	12-AUG-1999;	9905-0148341.
PR	13-AUG-1999;	9905-0148565.
PR	13-AUG-1999;	9905-0148684.
PR	16-AUG-1999;	9905-0149368.
PR	17-AUG-1999;	9905-0149175.
PR	18-AUG-1999;	9905-0149426.
PR	20-AUG-1999;	9905-0149722.
PR	20-AUG-1999;	9905-0149723.
PR	20-AUG-1999;	9905-0149929.
PR	23-AUG-1999;	9905-0149902.
PR	23-AUG-1999;	9905-0149930.
PR	25-AUG-1999;	9905-0150566.
PR	26-AUG-1999;	9905-0150884.
PR	27-AUG-1999;	9905-0151065.
PR	27-AUG-1999;	9905-0151066.
PR	27-AUG-1999;	9905-0151080.
PR	30-AUG-1999;	9905-0151303.
PR	31-AUG-1999;	9905-0151438.
PR	01-SEP-1999;	9905-0151930.
PR	07-SEP-1999;	9905-0152363.
PR	10-SEP-1999;	9905-0153070.
PR	13-SEP-1999;	9905-0153758.
PR	15-SEP-1999;	9905-0154018.
PR	16-SEP-1999;	9905-0154039.
PR	20-SEP-1999;	9905-0154779.
PR	22-SEP-1999;	9905-0155139.
PR	23-SEP-1999;	9905-0155486.
PR	24-SEP-1999;	9905-0155659.
PR	28-SEP-1999;	9905-0156458.
PR	29-SEP-1999;	9905-0156596.
PR	04-OCT-1999;	9905-0157117.
PR	05-OCT-1999;	9905-0157753.
PR	06-OCT-1999;	9905-0157865.
PR	07-OCT-1999;	9905-0158029.
PR	08-OCT-1999;	9905-0158222.
PR	12-OCT-1999;	9905-0158369.
PR	13-OCT-1999;	9905-0159293.
PR	13-OCT-1999;	9905-0159294.
PR	13-OCT-1999;	9905-0159295.
PR	14-OCT-1999;	9905-0159329.
PR	14-OCT-1999;	9905-0159330.
PR	14-OCT-1999;	9905-0159331.
PR	14-OCT-1999;	9905-0159637.
PR	14-OCT-1999;	9905-0159638.
PR	18-OCT-1999;	9905-0159584.
PR	21-OCT-1999;	9905-0160741.
PR	21-OCT-1999;	9905-0160767.
PR	21-OCT-1999;	9905-0160768.
PR	21-OCT-1999;	9905-0160770.
PR	21-OCT-1999;	9905-0160814.
PR	21-OCT-1999;	9905-0160815.
PR	22-OCT-1999;	9905-0160980.
PR	22-OCT-1999;	9905-0160981.
PR	22-OCT-1999;	9905-0160989.
PR	25-OCT-1999;	9905-0161404.
PR	25-OCT-1999;	9905-0161405.
PR	25-OCT-1999;	9905-0161406.
PR	26-OCT-1999;	9905-0161359.
PR	26-OCT-1999;	9905-0161360.
PR	26-OCT-1999;	9905-0161361.
PR	28-OCT-1999;	9905-0161920.
PR	28-OCT-1999;	9905-0161992.
PR	28-OCT-1999;	9905-0161993.
PR	29-OCT-1999;	9905-0162142.

DB	70	SALVAAAAA	79	
		RESULT 9		
		AAAG19453		
ID	AAAG19453	standard; Protein: 384 AA.		
XX	AAAG19453;			
XX				
DT	17-OCT-2000	(first entry)		
XX				
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 21261.			
XX				
KW	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter termination sequence.			
KW				
XX				
OS	Arabidopsis thaliana.			
XX	EP1033405-A2.			
PN				
XX				
PD	06-SEP-2000.			
XX				
PF	25-FEB-2000; 2000EP-0301439.			
XX				
PR	25-FEB-1999;	99US-0121825.		
PR	05-MAR-1999;	99US-0123180.		
PR	09-MAR-1999;	99US-0123548.		
PR	23-MAR-1999;	99US-0125788.		
PR	25-MAR-1999;	99US-0126264.		
PR	29-MAR-1999;	99US-0126785.		
PR	01-APR-1999;	99US-0127462.		
PR	06-APR-1999;	99US-0128234.		
PR	08-APR-1999;	99US-0128714.		
PR	16-APR-1999;	99US-0129845.		
PR	19-APR-1999;	99US-0130077.		
PR	21-APR-1999;	99US-0130449.		
PR	23-APR-1999;	99US-0130510.		
PR	23-APR-1999;	99US-0130891.		
PR	30-APR-1999;	99US-0131449.		
PR	30-APR-1999;	99US-0132048.		
PR	30-APR-1999;	99US-0132407.		
PR	04-MAY-1999;	99US-0132484.		
PR	05-MAY-1999;	99US-0132485.		
PR	06-MAY-1999;	99US-0132486.		
PR	06-MAY-1999;	99US-0132487.		
PR	07-MAY-1999;	99US-0132863.		
PR	11-MAY-1999;	99US-0134256.		
PR	14-MAY-1999;	99US-0134218.		
PR	14-MAY-1999;	99US-0134219.		
PR	14-MAY-1999;	99US-0134221.		
PR	14-MAY-1999;	99US-0134370.		
PR	18-MAY-1999;	99US-0134766.		
PR	19-MAY-1999;	99US-0134941.		
PR	20-MAY-1999;	99US-0135124.		
PR	21-MAY-1999;	99US-0135353.		
PR	24-MAY-1999;	99US-0135629.		
PR	25-MAY-1999;	99US-0136021.		
PR	27-MAY-1999;	99US-0136392.		
PR	28-MAY-1999;	99US-0136782.		
PR	01-JUN-1999;	99US-0137222.		
PR	03-JUN-1999;	99US-0137528.		
PR	04-JUN-1999;	99US-0137502.		
PR	07-JUN-1999;	99US-0137724.		
PR	08-JUN-1999;	99US-0138094.		
PR	10-JUN-1999;	99US-0138540.		
PR	10-JUN-1999;	99US-0138847.		
PR	14-JUN-1999;	99US-0139119.		
PR	16-JUN-1999;	99US-0139452.		
PR	16-JUN-1999;	99US-0139453.		
PR	17-JUN-1999;	99US-0139492.		
PR	18-JUN-1999;	99US-0139454.		
PR	18-JUN-1999;	99US-0139455.		

PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145216.
PR 27-JUL-1999; 99US-0145213.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160989.
PR 21-OCT-1999; 99US-0161040.
PR 21-OCT-1999; 99US-0161045.
PR 21-OCT-1999; 99US-0161081.
PR 21-OCT-1999; 99US-0161084.
PR 21-OCT-1999; 99US-0161085.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.9%; Score 10; DB 21; Length 384;
Best Local Similarity 100.0%; Pred. No. 0.84;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SALAVAAAA 19
Db 70 SALAVAAAA 79

RESULT 10
AACG9848
ID AACG9848 standard; Protein; 387 AA.
XX

AC AAG49848;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63104.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

```

PR 26-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151080.
PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155658.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161820.
PR 28-OCT-1999: 99US-0161992.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.

```

Query Match 2.9%; Score 10; DB 21; Length 387;

Best Local Similarity 100.0%; Pred. No. 0.85; Mismatches 0; Indels 0; Caps 0;

```

QY 10 SALAVAAAAA 19
   |||||
Db 70 SALAVAAAAA 79

```

RESULT 11

ID ABB62419 standard; protein; 337 AA.

AC ABB62419;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 14049.

KW Drosophila: developmental biology; cell signalling; insecticide;

```

KM pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL06522.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 14049; 21pp + Sequence Listing; English.
XX
PS The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL16176-ABL16175) and the encoded proteins
XX (ABB72072).
XX
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 337 AA;

```

Query Match 2.6%; Score 9; DB 22; Length 337;

Best Local Similarity 100.0%; Pred. No. 6.4; Mismatches 9; Caps 0;

```

QY 12 LAVAAAAAA 20
   |||||
Db 238 LAVAAAAAA 246

```

RESULT 12

ID AAB72380 standard; protein; 337 AA.

AC AAB72380;

DT 22-MAY-2001 (first entry)

DE Hairy protein amino acid sequence.

KW Cardiac helix-loop-helix factor; CHF; transcription factor;

KW cardiomyocyte; proliferation; myocarditis; myocardial infarction;

KW cardiomyocyte regeneration; angiogenesis inhibitor; differentiation;

KW smooth muscle cell growth; heart disease; hairy.

OS Drosophila sp.

PN WO200112126-A2.

PD 22-FEB-2001.

PF 11-AUG-2000; 2000WO-US21858.

PR 13-AUG-1999; 99US-0148974.

XX (HARD) HARVARD COLLEGE.
 PA (LEEM/) LEE M.
 XX
 PI Lee M, Chin MT;
 XX
 DR WPI; 2001-211119/21.
 XX
 PT Novel cardiovascular helix-loop-helix factor polypeptide and
 PT polynucleotide useful for regenerating heart tissue and promoting
 PT smooth muscle differentiation to treat heart disease or heart injury -
 XX
 PS Disclosure: Fig 1: 46pp: English.
 XX
 CC This invention relates to human cardiovascular helix-loop-helix factor
 CC (CHF) polypeptides. CHF is a transcription factor expressed in
 CC cardiovascular tissue, which contains a basic helix-loop-helix domain.
 CC CHF-1 shares structural homology with hairy, a Drosophila protein
 CC essential for the development of the peripheral nervous system. The CHF
 CC polypeptide is useful for expanding an endogenous population of
 CC proliferative cardiomyocytes in a heart tissue, promoting proliferation
 CC of a cardiomyocyte in a mammal having myocarditis or that which has
 CC suffered myocardial infarction and for regenerating cardiomyocytes in
 CC vivo. DNA encoding a CHF-1 polypeptide is also useful for promoting
 CC proliferation of a cardiomyocyte, regenerating cardiomyocytes in vitro,
 CC inhibiting angiogenesis in a tissue, inducing differentiation of smooth
 CC muscle cells in a mammalian tissue, preferably venous tissue, inducing
 CC growth of smooth muscle cells in a vein explant, promoting smooth muscle
 CC cell regeneration in an injured or diseased vascular tissue (venous or
 CC arterial) and for reducing vein graft stenosis in a mammal by contacting
 CC the tissue ex vivo with the DNA prior to implantation of the tissue into
 CC an artery of the mammal. The present sequence represents the hairy
 CC protein from Drosophila, with which CHF-1 shares homology.
 XX

SQ Sequence 337 AA:

Query Match 2.6%; Score 9; DB 22; Length 337;
 Best Local Similarity 100.0%; Pred. No. 6.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LAVAAAAA 20
 |||||
 Db 238 LAVAAAAA 246

RESULT 13

AAB99727
 ID AAB99727 standard; Protein: 350 AA.

XX AAB99727;
 XX

DT 10-SEP-2001 (first entry)

DE Streptomyces sp. CL190 mevalonate pathway orfB protein SEQ ID NO:9.

XX Streptomyces sp. CL190; mevalonate pathway; actinomycete; cardiant;
 KM isoprenoid compound; osteopathic; cytosolic; ubiquitome; vitamin K2;
 KM carotenoid; heart disease; osteoporosis; cancer; drug; health food.
 XX

OS Streptomyces sp. CL190.

XX WO200142476-A1.
 XX

PD 14-JUN-2001.

XX 06-DEC-2000; 2000WO-JP08620.
 XX

PR 08-DEC-1999; 99JP-0348375.
 XX

PA (SETO/) SETO H.
 PA (KUZU/) KUZUYAMA T.

PI Seto H, Kuzuyama T, Takahashi S, Takagi M;

XX WPI: 2001-381696/40.
 DR N-PSDB; AAA44045.
 XX
 PT Actinomycetes-originated genes of enzymes participating in mevalonate
 PT pathway, applicable in producing e.g. ubiquitome, vitamin K2 and
 PT carotenoids for treatment of heart diseases, osteoporosis and cancer in
 PT drug and health food -
 XX
 PS Claim 6; Page 60-62; 75pp: Japanese.
 XX

CC The sequence given in AAA44043 represents a DNA sequence isolated from
 CC Streptomyces sp. CL190, containing a 6798 base pairs (S1), which encodes
 CC the whole enzyme necessary for functioning the mevalonate pathway. The
 CC sequence encodes protein sequences, designated orfA to E and hmgt, which
 CC are used in the mevalonate pathway. The proteins and polynucleotide
 CC sequences encoding them have cardiant, osteopathic and cytosolic
 CC activities. The genes are applicable in producing e.g. ubiquitome,
 CC vitamin K2 and carotenoids which can be used in the treatment of heart
 CC diseases, osteoporosis and cancer in drugs and health foods. The present
 CC sequence represents the orfB protein from the present invention.

SQ Sequence 350 AA:

Query Match 2.6%; Score 9; DB 22; Length 350;
 Best Local Similarity 100.0%; Pred. No. 6.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ALAVAAAA 19
 |||||
 Db 130 ALAVAAAA 138

RESULT 14

AAB69366
 ID AAB69366 standard; Protein: 1068 AA.

XX AAB69366;
 XX

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 34890.

XX Drosophila: developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 XX

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.
 XX

PR 23-MAR-2000; 2000US-191637P.
 XX

PR 11-JUL-2000; 2000US-0614150.
 XX

PA (PEKE) PE CORP NY.
 XX

XX Venter JC, Adams M, Li PWD, Myers EW;
 PI

DR WPI: 2001-656860/75.
 DR N-PSDB; ABL13469.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX

PS Disclosure: SEQ ID NO 34890; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1068 AA;

Query Match 2.6%; Score 9; DB 22; Length 1068;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LAVAAAAA 20
| | | | | | | | | |
DB 340 LAVAAAAA 348

RESULT 15

ABB63936

ID ABB63936 standard; Protein: 1307 AA.

AC ABB63936;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 18600.

KW Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL08039.

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

PS Disclosure; SEQ ID NO 18600; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 1307 AA;

Query Match

Best Local Similarity 2.6%; Score 9; DB 22; Length 1307;
100.0%; Pred. No. 22;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 275 SSAPASSS 283
| | | | | | | | | |
DB 379 SSAPASSS 387

Search completed: July 11, 2003, 11:04:09
Job time : 72 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:01:55 ; Search time 41 Seconds
(without alignments)
818.315 Million cell updates/sec

Title: US-09-654-652a-3

Sequence: 1 MNIKRTAVKSAALVAIAAAAAA.....AKGAKVPMNGHKRYRVNFEH 349

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	100.0	349	2 A44507	licheninase (EC 3.
2	10	2.9	348	2 T52635	mitogen-activated
3	10	2.9	348	2 T51340	mitogen-activated
4	10	2.9	621	1 S59632	endo-1,4-beta-xyla
5	9	2.6	184	2 B84259	hypothetical prote
6	9	2.6	337	2 S06956	segmentation prote
7	9	2.6	439	2 A72599	hypothetical prote
8	9	2.6	537	1 A46600	methylmalonate-sem
9	9	2.6	776	2 T29064	hyaluronate lyase
10	9	2.6	918	1 I48719	protein kinase C (
11	9	2.6	1475	2 S42718	nuclear pore compl
12	8	2.3	37	1 FDFL3W	antifreeze protein
13	8	2.3	45	2 PNO589	tyrosine 3-monooxy
14	8	2.3	45	2 PNO590	tyrosine 3-monooxy
15	8	2.3	45	2 PNO591	tyrosine 3-monooxy
16	8	2.3	45	2 PNO592	tyrosine 3-monooxy
17	8	2.3	45	2 PNO593	tyrosine 3-monooxy
18	8	2.3	61	2 G41476	probable antigen 7
19	8	2.3	82	1 FDFLAW	antifreeze protein
20	8	2.3	82	1 S02126	antifreeze protein
21	8	2.3	82	2 JS0706	antifreeze protein
22	8	2.3	82	2 A05161	antifreeze protein
23	8	2.3	82	2 JS0705	antifreeze protein
24	8	2.3	82	2 I51125	antifreeze protein
25	8	2.3	158	2 T34030	hypothetical prote
26	8	2.3	160	2 H86978	6,7-dimethyl-8-r
27	8	2.3	219	2 I51382	achate-scutle homo
28	8	2.3	231	2 E81215	thiol-disulfide in
29	8	2.3	284	2 T06141	probable receptor-

30	8	2.3	291	2 T51668	myb-related transc
31	8	2.3	306	2 A88040	protein F47F6.1 [i
32	8	2.3	317	2 C70874	hypothetical prote
33	8	2.3	320	2 B87135	conserved hypotet
34	8	2.3	339	2 T26328	hypothetical prote
35	8	2.3	368	2 G83463	probable methyltra
36	8	2.3	368	2 T46615	chemotaxis proteob
37	8	2.3	374	2 T03875	probable homeobox
38	8	2.3	375	2 T03874	probable homeobox
39	8	2.3	378	2 T06512	DNA-binding protei
40	8	2.3	381	2 S29560	fructose-bisphosph
41	8	2.3	392	2 B48423	homeotic protein e
42	8	2.3	401	2 A48423	engrailed homeodom
43	8	2.3	407	2 T06408	probable fructose-
44	8	2.3	414	2 A86229	hypothetical prote
45	8	2.3	424	2 S09884	hypothetical prote

ALIGNMENTS

RESULT 1
A44507
licheninase (EC 3.2.1.75) - Fibrobacter succinogenes
C:Species: Fibrobacter succinogenes
C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 08-Oct-1999
C:Accession: A44507
R:Reather, R.M.; Erfile, J.D.
J. Bacteriol. 172, 3837-3841, 1990
A:Title: DNA sequence of a Fibrobacter succinogenes mixed-linkage beta-glucanase (1,3
A:Reference number: A44507; MUID:90299807; PMID:2193918
A:Accession: A44507
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <TEA>
A:Cross-references: EMBL:M33676; NID:g148575; PIDN:AAA24896.1; PID:g148576
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 100.0%; Score 349; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MNIIKRTAVKSAALVAIAAAALTTNVSADPSGAEIYTLBEVOYGFRRKMAAAGTYS	60
DB	1	MNIIKRTAVKSAALVAIAAAALTTNVSADPSGAEIYTLBEVOYGFRRKMAAAGTYS	60
QY	61	SMFLYONGSEIADGRVWEVDIEVLGKNPGSFQSNITTGKAGAKTSEKHHAASPAADQA	120
DB	61	SMFLYONGSEIADGRVWEVDIEVLGKNPGSFQSNITTGKAGAKTSEKHHAASPAADQA	120
QY	121	FHTYGLMTFNRYRWYVDGQEVKRTBEGQVSNLTGTQGLRFLNLMSSSAWVQFDESKL	180
DB	121	FHTYGLMTFNRYRWYVDGQEVKRTBEGQVSNLTGTQGLRFLNLMSSSAWVQFDESKL	180
QY	181	PLFOFINMWVLYVYKTPGQGGSGDFTLDMTDNFTDPGSRWKGMDTFDGNRYDLTDKNI	240
DB	181	PLFOFINMWVLYVYKTPGQGGSGDFTLDMTDNFTDPGSRWKGMDTFDGNRYDLTDKNI	240
QY	241	YSRDGMILIALTRKGOESFNGQVPRDEPAPOSSSSAPASSSSVPASSSSVPAASSSAFV	300
DB	241	YSRDGMILIALTRKGOESFNGQVPRDEPAPOSSSSAPASSSSVPASSSSVPAASSSAFV	300
QY	301	PPSSSSATNIAHGMRTTPAVAKKRNILVNAKGAQVPMNGHKRYRVNFEH	349
DB	301	PPSSSSATNIAHGMRTTPAVAKKRNILVNAKGAQVPMNGHKRYRVNFEH	349

RESULT 2
T52635
mitogen-activated protein kinase (EC 2.7.1.1) alpha [imported] - Arabidopsis t
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 03-Nov-2000
C:Accession: T52635

R:Hamal, A.; Jouanin, S.; Lepoint, S.; Kreiss, M.; Henry, Y.
Plant Sci. 140, 41-52, 1999
A:Title: Molecular characterization and expression of an Arabidopsis thaliana L. MAP kin
A:Reference number: Z26147
A:Accession: T52635
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-348 <HAM>
A:Cross-references: EMBL:Y07694; PDB:CAA68958.1
A:Experimental source: cultivar Columbia; seedling
C:Genetics:
A:Gene: MAP2Kalpha
A:Map position: 3
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase

Query Match 2.9%; Score 10; DB 2; Length 348;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 273 SSSAPASS 282
DB 46 SSSAPASS 55

RESULT 3
T51340
Mitogen-activated protein kinase kinase (EC 2.7.1.-) 5 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 08-Sep-2000
C:Accession: T51340
R:Ichimura, K.; Mizoguchi, T.; Hayashida, N.; Seki, M.; Shinzaki, K.
DNA Res. 5, 341-348, 1998
A:Title: Molecular cloning and characterization of three cDNAs encoding putative mitogen
A:Reference number: Z25272; MUID:9916228; PMID:10048483
A:Accession: T51340
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-348 <ICH>
A:Cross-references: EMBL:AB015316; PDB:BAA28831.1
C:Genetics:
A:Gene: ATKMS
C:Function:
A:Description: (EC 2.7.1.-); mitogen-activated protein kinase kinase [validated, MUID:99
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: phosphotransferase; protein kinase

Query Match 2.9%; Score 10; DB 2; Length 348;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 273 SSSAPASS 282
DB 46 SSSAPASS 55

RESULT 4
S59632
endo-1,4-beta-xylosylase (EC 3.2.1.8) B precursor - Cellvibrio mixtus
C:Species: Cellvibrio mixtus
C:Date: 15-Feb-1996 #sequence_revision 11-Apr-1997 #text_change 18-Jun-1999
C:Accession: S59632; S52742
R:Millward-Sadler, S.J.; Davidson, K.; Hazlewood, G.P.; Black, G.W.; Gilbert, H.J.; Clar
Biochem. J. 312, 39-48, 1995
A:Title: Novel cellulose-binding domains, NodB homologues and conserved modular architec
A:Reference number: S59631; MUID:96077124; PMID:7492333
A:Accession: S59632
A:Molecule type: DNA
A:Residues: 1-621 <MIL>
A:Cross-references: EMBL:Z48926; NID:g757808; PDB:CAA88762.1; PID:g757809
C:Genetics:

A:Gene: xynB
C:Function:
A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xyans
A:Pathway: xylan degradation
C:Superfamily: Pseudomonas endo-1,4-beta-xylosylase F; Streptomyces endo-1,4-beta-xylan
C:Keywords: glycosylase; hydrolase; polysaccharide degradation
F:1-19/Domain: signal sequence #status predicted <Sig>
F:20-621/Product: endo-1,4-beta-xylosylase B #status predicted <MAT>
F:302-615/Domain: Streptomyces endo-1,4-beta-xylosylase A homology <SXY>
F:403,516/Active site: Glu #status predicted

Query Match 2.9%; Score 10; DB 1; Length 621;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 274 SSSAPASS 283
DB 89 SSSAPASS 98

RESULT 5
B84259
hypothetical protein Vng1026h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: B84259
R:Ni, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky
; Leitauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.;
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: B84259
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-184 <STO>
A:Cross-references: GB:AE004437; NID:g10580580; PDB:AG19438.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG1026H

Query Match 2.6%; Score 9; DB 2; Length 184;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 LAVAAAAA 20
DB 137 LAVAAAAA 145

RESULT 6
S06956
segmentation protein hairy - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 21-Jul-2000
C:Accession: S06956; S06700
R:Rushlow, C.A.; Hogan, A.; Pichin, S.M.; Howe, K.M.; Lardelli, M.; Ish-Horowitz, D.
EMBO J. 8, 3095-3103, 1989
A:Title: The Drosophila hairy protein acts in both segmentation and bristle patternin
A:Reference number: S06956; MUID:90059896; PMID:2479541
A:Accession: S06956
A:Molecule type: DNA
A:Residues: 1-337 <RUS>
A:Cross-references: GB:X15904; GB:S63792; GB:X16632; NID:g8048; PDB:CAA34018.1; PID:
A:Note: 252-ser was also found
C:Genetics:
A:Gene: hairy
A:Cross-references: FlyBase:FBgn0001168
A:Map position: 66D
A:Introns: 33/3; 65/3
C:Keywords: DNA binding; transcription regulation

Query Match 2.6%; Score 9; DB 2; Length 337;

Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LAVAAAAA 20
|||||
Db 238 LAVAAAAA 246

RESULT 7

A72599

hypothetical protein APE1257 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: A72599

R:Kawabuchi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatakeyama, Y.; Jin-no, K.; Takai,

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: A72599

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-439 <RAW>

A:Cross-references: DDBJ:AP000061; NID:95104821; PIDN:BAA80247.1; PID:01044033; PID:9510

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1257

Query Match

2.6%; Score 9; DB 2; Length 439;
Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 SALAVAAA 18
|||||
Db 398 SALAVAAA 406

RESULT 8

A46600

methylmalonate-semialdehyde dehydrogenase (acylating) (EC 1.2.1.27) precursor - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-May-2000

C:Accession: A46600

R:Deichate, I.; Berthiaume, L.; Peseckis, S.M.; Patton, W.F.; Resh, M.D.

J. Biol. Chem. 268, 13738-13747, 1993

A>Title: Novel use of an lodo-myristyl-CoA analog identifies a semialdehyde dehydrogenase

A:Reference number: A46600; MUID:93293905; PMID:8514806

A:Accession: A46600

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-537 <DEI>

A:Cross-references: GB:108643; NID:q289441; PIDN:AAA30650.1; PID:q289442

C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

C:Keywords: coenzyme A; mitochondrion; oxidoreductase

F:80-338/Domain: aldehyde dehydrogenase homology <ALDD>

F:319/Active site: Cys #status predicted

Query Match

2.6%; Score 9; DB 1; Length 537;
Best Local Similarity 100.0%; Pred. No. 3.2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 AVAAAAAL 21
|||||
Db 5 AVAAAAAL 13

RESULT 9

T29064

hyaluronate lyase homolog - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000

C:Accession: T29064

R:Redenbach, M.; Kleser, H.M.; Denapalte, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopa

Mol. Microbiol. 21, 77-96, 1996

A>Title: A set of ordered cosmid and a detailed genetic and physical map for the 8 M

A:Reference number: Z20555; MUID:97000351; PMID:8843436

A:Accession: T29064

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-776 <RED>

A:Cross-references: EMBL:AL031124; NID:e1312893; PID:e1312908; PIDN:CAA19982.1

C:Genetics:

A>Note: SCIC2.15

Query Match

2.6%; Score 9; DB 2; Length 776;
Best Local Similarity 100.0%; Pred. No. 4.4;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 AVAAAAALTT 23
|||||
Db 18 AVAAAAALTT 26

RESULT 10

I48719

protein kinase C (EC 2.7.1.-) mu precursor - mouse

N:Alternate names: protein kinase D

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 11-Jun-1999

C:Accession: I48719

R:Valverde, A.M.; Sinnett-Smith, J.; Van Lint, J.; Rozengurt, E.

Proc. Natl. Acad. Sci. U.S.A. 91, 8572-8576, 1994

A>Title: Molecular cloning and characterization of protein kinase D: a target for dia

A:Reference number: I48719; MUID:94359973; PMID:8078925

A:Accession: I48719

A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-918 <RES>

A:Cross-references: EMBL:Z34524; NID:9520877; PIDN:CAA84283.1; PID:9520878

C:Function:

A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threo

sely related human enzyme, this protein is reported to bind phorbol esters

C:Superfamily: protein kinase C mu; protein kinase C zinc-binding repeat homology; pr

C:Keywords: ATP; autophosphorylation; duplication; phorbol ester binding; phospholipid

F:1-25/Domain: signal sequence #status predicted <SIG>

F:26-918/Product: protein kinase mu #status predicted <MAT>

F:145-194/Domain: protein kinase C zinc-binding repeat homology <KZ1>

F:277-326/Domain: protein kinase C zinc-binding repeat homology <KZ2>

F:387-845/Domain: protein kinase homology <KIN>

F:595-603/Region: protein kinase ATP-binding motif

F:618-636,712,714/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match

2.6%; Score 9; DB 1; Length 918;
Best Local Similarity 100.0%; Pred. No. 5.1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 AVAAAAAL 21
|||||
Db 19 AVAAAAAL 27

RESULT 11

S42718

nuclear pore complex protein nup153 - human

C:Species: Homo sapiens (man)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999

C:Accession: S42718; S37477

R:McMorrow, I.; Bastos, R.; Horton, H.; Burke, B.

Biochim. Biophys. Acta 1217, 219-223, 1994

A>Title: Sequence analysis of a cDNA encoding a human nuclear pore complex protein, h

A:Reference number: S42718; MUID:94154002; PMID:8110839

A:Accession: S42718

A:Molecule type: mRNA

A:Residues: 1-1475 <KCM>

A:Cross-references: EMBL:Z25535; NID:9406224; PIDN:CAA80982.1; PID:9406225

Query Match 2.3%; Score 8; DB 2; Length 1475;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 282 SSVPASSSS 290
|||||
Db 823 SSVPASSSS 831

RESULT 12

PDL3W

antifreeze protein 3 - winter flounder

C:Species: Pseudopleuronectes americanus (winter flounder)

C>Date: 01-Sep-1981 #sequence_revision 01-Sep-1981 #text_change 23-Aug-1996

C:Accession: A03192

R:Devries, A.L.; Lin, Y

Biochim. Biophys. Acta 495, 388-392, 1977

A:Title: Structure of a peptide antifreeze and mechanism of adsorption to ice.

A:Reference number: A03192; MUID:78060969; PMID:588591

A:Accession: A03192

A:Molecule type: protein

A:Residues: 1-37 <DEV>

C:Superfamily: antifreeze protein

C:Keywords: antifreeze

Query Match 2.3%; Score 8; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 AAAAAALT 22
|||||
Db 6 AAAAAALT 13

RESULT 13

PN0589

tyrosine 3-monooxygenase (EC 1.14.16.2) type 1 and 2 - chimpanzee (fragment)

N:Alternate names: oxygen oxidoreductase; tetrahydropteridine; tyrosine hydroxylase

C:Species: Pan troglodytes (chimpanzee)

C>Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 31-Mar-2000

C:Accession: PN0589

R:Richinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.

Biochem. Biophys. Res. Commun. 195, 158-165, 1993

A:Title: Increased heterogeneity of tyrosine hydroxylase in humans.

A:Reference number: PN0575; MUID:93371398; PMID:7689834

A:Accession: PN0589

A:Molecule type: genomic RNA

A:Residues: 1-45 <ICH>

A:Cross-references: GB:L14791

A:Experimental source: lymphocytes of peripheral blood

C:Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine biosynthesis

C:Superfamily: phenylalanine 4-monooxygenase

C:Keywords: bioperin; monoxygenase; oxidoreductase

Query Match 2.3%; Score 8; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 AVAAAAA 20
|||||
Db 22 AVAAAAA 29

RESULT 14

PN0590

tyrosine 3-monooxygenase (EC 1.14.16.2) type 1 and 2 - gorilla (fragment)

N:Alternate names: oxygen oxidoreductase; tetrahydropteridine; tyrosine hydroxylase

C:Species: Gorilla gorilla (gorilla)

C>Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 31-Mar-2000

C:Accession: PN0590

R:Richinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.

Biochem. Biophys. Res. Commun. 195, 158-165, 1993

A:Title: Increased heterogeneity of tyrosine hydroxylase in humans.

A:Reference number: PN0575; MUID:93371398; PMID:7689834

A:Accession: PN0590

A:Molecule type: genomic RNA

A:Residues: 1-45 <ICH>

A:Cross-references: GB:L14797

A:Experimental source: lymphocytes of peripheral blood

C:Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine biosynthesis

C:Superfamily: phenylalanine 4-monooxygenase

C:Keywords: bioperin; monoxygenase; oxidoreductase

Query Match 2.3%; Score 8; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 AVAAAAA 20
|||||
Db 22 AVAAAAA 29

RESULT 15

PN0591

tyrosine 3-monooxygenase (EC 1.14.16.2) type 1 and 2 - orangutan (fragment)

N:Alternate names: oxygen oxidoreductase; tetrahydropteridine; tyrosine hydroxylase

C:Species: Pongo pygmaeus (orangutan)

C>Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 31-Mar-2000

C:Accession: PN0591

R:Richinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.

Biochem. Biophys. Res. Commun. 195, 158-165, 1993

A:Title: Increased heterogeneity of tyrosine hydroxylase in humans.

A:Reference number: PN0575; MUID:93371398; PMID:7689834

A:Accession: PN0591

A:Molecule type: genomic RNA

A:Residues: 1-45 <ICH>

A:Cross-references: GB:L14800

A:Experimental source: lymphocytes of peripheral blood

C:Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine biosynthesis

C:Superfamily: phenylalanine 4-monooxygenase

C:Keywords: bioperin; monoxygenase; oxidoreductase

Query Match 2.3%; Score 8; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 AVAAAAA 20
|||||
Db 22 AVAAAAA 29

Search completed: July 11, 2003, 11:06:56
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 10:54:40 ; Search time 23 Seconds

(without alignments)
629.358 Million cell updates/sec

Title: US-09-654-652a-3

Perfect score: 349
Sequence: 1 MNKKTAVKASALAAVAAAAA.....AKGAKVNPNGHKRYRVNFEH 349

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	100.0	349	1 GUB_FIBSU	P17889 fibrobacter
2	10	2.9	241	1 TIRA_MOUSE	Q991y1 mus musculus
3	9	2.6	337	1 HAIR_DROME	P14003 drosophila
4	9	2.6	537	1 MMSA_BOVIN	Q07336 bos taurus
5	9	2.6	918	1 KPCM_MOUSE	Q62101 mus musculus
6	9	2.6	1475	1 N153_HUMAN	P49790 homo sapien
7	8	2.3	37	1 ANP3_PSEAM	P02733 pseudopleur
8	8	2.3	82	1 ANPA_PSEAM	P04002 pseudopleur
9	8	2.3	133	1 RS16_RHIL	Q98670 rhizobium
10	8	2.3	160	1 RISB_MYCLE	Q9CCP3 mycobacteri
11	8	2.3	268	1 ZEST_DROMA	Q24597 drosophila
12	8	2.3	286	1 ALI4_ASFRU	Q60024 aspergillus
13	8	2.3	317	1 CYSD_RHIME	P56892 rhizobium m
14	8	2.3	317	1 YE80_MYCTU	P71161 mycobacteri
15	8	2.3	368	1 CHEB_PSEAE	O87125 pseudomonas
16	8	2.3	387	1 SOX1_HUMAN	O00570 homo sapien
17	8	2.3	391	1 SOX1_MOUSE	P53783 mus musculu
18	8	2.3	392	1 HME1_HUMAN	Q05525 homo sapien
19	8	2.3	401	1 HME1_MOUSE	P09065 mus musculu
20	8	2.3	407	1 FLIP_PEA	P46770 pisum sativ
21	8	2.3	424	1 ULB7_HOMVA	P16770 human cytom
22	8	2.3	428	1 FXB2_MOUSE	Q64733 mus musculu
23	8	2.3	429	1 OGRA_MYCTU	O10387 mycobacteri
24	8	2.3	440	1 DCO_DROME	O76324 drosophila
25	8	2.3	459	1 IP3K_RAT	P17105 rattus norv
26	8	2.3	461	1 IP3K_HUMAN	P23677 homo sapien
27	8	2.3	462	1 A2AC_HUMAN	P18825 homo sapien
28	8	2.3	470	1 ESCA_DROME	P25932 drosophila
29	8	2.3	483	1 ELAV_DROVI	P16914 drosophila
30	8	2.3	519	1 ELAV_DROVI	P23441 drosophila
31	8	2.3	528	1 TY3H_HUMAN	P07101 homo sapien
32	8	2.3	542	1 CH12_RHIL	P29027 rhizopus ol
33	8	2.3	559	1 PHF1_MOUSE	Q921b8 mus musculu

ALIGNMENTS

```

RESULT 1
GUB_FIBSU
ID GUB_FIBSU          STANDARD:      PRT:      349 AA.
AC P17889:
DT 01-NOV-1990 (rel. 16, Created)
DT 01-NOV-1990 (rel. 16, Last sequence update)
DT 15-JUL-1999 (rel. 38, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Mixed linkage beta-
DE glucanase) (Lichenase).
OS Fibrobacter succinogenes (Bacteroides succinogenes).
OC Bacteria: Fibrobacter: Acidobacteria group: Fibrobacter group:
OC Fibrobacter.
OX NCBI_TaxID=833:
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-57.
RC STRAIN=Isolate S85;
RX MEDLINE=90299807; PubMed=2193918;
RA Teather R.M., Erfle J.D.;
RT *DNA sequence of a Fibrobacter succinogenes mixed-linkage
RT beta-glucanase (1,3-1,4-beta-D-glucan 4-glucanohydrolase) gene.
RL J. Bacteriol. 172:3837-3841(1990).
CC -I- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -I- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M3676; AAA24396.1; -.
CC PIR: A44507; A44507.
CC HSSP: P23904; IAIK.
CC InterPro: IPR000757; Glyco_hydro_16.
CC Pfam: PRO0722; Glyco_hydro_16; 1.
CC PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
CC KEGG: Glycolase; Glycosidase; Signal; Repeat.
CC SIGNAL
CC FT 1 28 349 FT
CC ACT_SITE 79 79
CC ACT_SITE 83 83
CC DOMAIN 271 307
CC REPEAT 271 277
CC REPEAT 278 284
CC REPEAT 285 291
CC REPEAT 292 298
CC REPEAT 301 307
CC SEQUENCE 349 AA; 37737 MW; 16DC4F5BDFEC578A CRC64;
Query Match 100.0%; Score 349; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 7.9e-311;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 MNIKTAVKALAVAAAAAALTTNVSADPSCAGELTLEBYOYKFEARBMKMAAGTYS 60
        |||||||
DB      1 MNIKTAVKALAVAAAAAALTTNVSADPSCAGELTLEBYOYKFEARBMKMAAGTYS 60
OY      61 SMELVONGSEIADGRPMVEVDIEVLKNGPSFQSNITTKAGAKQKTEKHHAVSPAADA 120
        |||||||
DB      61 SMELVONGSEIADGRPMVEVDIEVLKNGPSFQSNITTKAGAKQKTEKHHAVSPAADA 120
OY      121 FHTYGLKMTPNVYRWVDDQEVARKTGGGVSNLTGQGLRPNLMSSESAAMVQGFESKL 180
        |||||||
DB      121 FHTYGLKMTPNVYRWVDDQEVARKTGGGVSNLTGQGLRPNLMSSESAAMVQGFESKL 180
OY      181 PLTFQFINWKKVYKYPGCGEGSDFTLDWTDNFDTPDGSRMCKGDMTFPGNRYDLTDKNI 240
        |||||||
DB      181 PLTFQFINWKKVYKYPGCGEGSDFTLDWTDNFDTPDGSRMCKGDMTFPGNRYDLTDKNI 240
OY      241 YSRDGMILIALRRKGSEFNGOVRDDEPAPQSSAPASSSVAPASSSVAPASSSAFV 300
        |||||||
DB      241 YSRDGMILIALRRKGSEFNGOVRDDEPAPQSSAPASSSVAPASSSVAPASSSAFV 300
OY      301 PSSSSATNAGHGMRTTPAVAKEHRLVNAKGAKVNPNGHRRVNFEEH 349
        |||||||
DB      301 PSSSSATNAGHGMRTTPAVAKEHRLVNAKGAKVNPNGHRRVNFEEH 349

```

RESULT 2

```

TIRA_MOUSE STANDARD; PRT; 241 AA.
ID TIRA_MOUSE
AC 099JY1; 0912W0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Toll-interleukin 1 receptor domain-containing adapter protein (TIR
DE domain-containing adapter protein) (MyD88 adapter-like protein)
DE (Adapter protein WYATL).
GN TIRAP OR MAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC STRAIN=C57BL/6;
RX MEDLINE=21417617; PubMed=11526399;
RA Horng T., Barton G.M., Medzhitov R.;
RT "TIRAP: an adapter molecule in the Toll signaling pathway.";
RN [2]
RA Nat. Immunol. 2:835-841(2001).
RN [3]
RC SEQUENCE FROM N.A.
RA KIRK P.B., Pereira J.P., Bazan F.;
RT "Characterization and structural analysis of TIR domain-containing
RT adaptor protein WYATL.";
RN [4]
RA Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [8]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [9]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [10]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [11]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [12]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [13]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [14]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [15]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [16]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [17]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [18]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [19]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [20]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [21]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [22]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [23]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [24]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [25]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [26]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [27]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [28]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [29]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [30]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [31]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [32]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [33]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [34]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [35]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [36]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [37]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [38]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [39]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [40]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [41]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [42]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [43]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [44]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [45]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [46]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [47]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [48]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [49]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [50]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [51]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [52]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [53]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [54]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [55]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [56]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [57]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [58]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [59]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [60]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [61]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [62]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [63]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [64]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [65]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [66]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [67]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [68]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [69]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [70]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [71]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [72]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [73]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [74]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [75]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [76]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [77]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [78]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [79]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [80]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [81]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [82]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [83]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [84]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [85]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [86]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [87]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [88]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [89]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [90]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [91]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [92]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [93]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [94]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [95]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [96]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [97]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [98]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [99]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [100]
RC SEQUENCE FROM N.A.
RA STRAUSBERG R.;
RT "Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AF378130; AAL05628.1; -
CC DR EMBL; AF410784; AAL05037.1; -
CC DR EMBL; BC005584; AAL05584.1; -
CC DR MGI; 2152213; TIRap.
CC DR InterPro: IPR000157; TIR_domain.
CC DR PROSITE: PS50104; TIR_1.
CC KW Immune response; Inflammatory response.
CC FT DOMAIN 104
CC FT POLY-SER.
CC FT CONFLICT 164 164 A -> V (IN REF. 1).
CC SQ SEQUENCE 241 AA; 26035 MW; C387D20229E12265 CRC64;

```

Query Match

```

Best local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

```

```

OY      279 ASSSVPASS 288
        |||||||
DB      2 ASSSVPASS 11

```

RESULT 3

```

HAIR_DROME STANDARD; PRT; 337 AA.
ID HAIR_DROME
AC P14003; 09VSN8;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hair protein.
GN H OR CG6494.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RC SEQUENCE FROM N.A.
RA STRAIN=Oregon-R;
RX MEDLINE=90059896; PubMed=2479541;
RA Rushlow C.A., Hogan A., Pterchin S.M., Howe K.M., Lardelli M.,
RA Ish-Horowicz D.;
RT "The Drosophila hair protein acts in both segmentation and bristle
RT patterning and shows homology to N-myc.";
RN [2]
RA EMBO J. 8:3095-3103(1989).
RN [3]
RC SEQUENCE FROM N.A.
RA STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abmayant A., An H.-J., Andrews-Ffannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari P., Brottier P.,
RA Borotova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Dwyer A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durkin K.J., Evangelista C.C., Ferraz C., Ferrelira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Galburt W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,

```

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., Moperson D.,
RA Merklov G., Mlshina N.V., Mobarry C., Morris J., Moshnefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Peltman K.A., Nixon K., Nusskern D.R., Pacleby J.M.,
RA Pelazolo M., Piltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*,"
RL Science 287:2185-2195(2000).
RN [3]
RP WRPW MOTIF.
RX MEDLINE:95094252; PubMed-8001118;
RA Paroush Z., Finley R.L. Jr., Kidd T., Wainwright S.M., Ingham P.W.,
RA Brent R., Ish-Horowicz D.,
RT "Groucho is required for *Drosophila* neurogenesis, segmentation, and
RT sex determination and interacts directly with hairy-related bHLH
RT proteins.";
RL Cell 79:805-815(1994).
CC -1- FUNCTION: PAIR-RULE PROTEIN THAT REGULATES EMBRYONIC SEGMENTATION
CC AND ADULT BRISTLE PATTERNING. TRANSCRIPTIONAL REPRESSOR OF GENES
CC THAT REQUIRE A BHLH PROTEIN FOR THEIR TRANSCRIPTION (EG. THE FUSHI
CC TARAU GENE).
CC -1- SUBUNIT: TRANSCRIPTION REPRESSION REQUIRES FORMATION OF A COMPLEX
CC WITH A CO-REPRESSOR PROTEIN (GROUCHO).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A
CC HELIX-INTERUPTING PROLINE) THAT BINDS TO THE N-BOX (CACNMG),
CC RATHER THAN THE CANONICAL E-BOX (CANMG).
CC -1- DOMAIN: THE CARBOXYL-TERMINAL WRPW MOTIF IS A TRANSCRIPTIONAL
CC REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO. A
CC TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY
CC HAIR-RELATED PROTEINS.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15904; CA334018.1; -;
DR EMBL: X15905; CA334019.1; -;
DR EMBL: AE003554; AAF50378.1; -;
DR PIR: S06956; S06956.
DR TRANSFAC: T00345; -;
DR FLYBASE: FBgn0001168; h.
DR InterPro: IPR001092; HLH_basic.
DR InterPro: IPR003650; Orange.
DR SMART: PF00010; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR SMART: SM00511; ORANGE; 1.
DR PROSITE: PS00038; HLH_1; 1.
DR PROSITE: PS50888; HLH_2; 1.
KW Nuclear protein; Developmental protein; Pair-rule protein;
KW DNA-binding; Transcription regulation; Repressor; Polymorphism.
FT DNA BIND 32 44
FT DOMAIN 149 89
FT DOMAIN 149 157
FT DOMAIN 222 237
FT DOMAIN 241 250
FT DOMAIN 334 337
FT VARIANT 292 292
S -> P.

SQL SEQUENCE 337 AA; 36995 MW; 6D2ECA7F2D56C0B CRC64;
Query Match 2.6%; Score 9; DB 1; Length 337;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 12 LAVAAAAA 20
DB 238 LAVAAAAA 246
RESULT 4
MMSA_BOVIN STANDARD; PRT; 537 AA.
ID MMSA_BOVIN
AC 007536;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial
DE precursor (EC 1.2.1.27) (MMSDH).
GN ALDH6A1 OR MMSDH.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE:93293905; PubMed-8514806;
RT Dechaete I., Berthiaume L., Pesceckis S.M., Patton W.F., Resh M.D.;
RT "Novel use of an lodo-myristyl-CoA analog identifies a semialdehyde
RT dehydrogenase in bovine liver.";
RL J. Biol. Chem. 268:13738-13747(1993).
CC -1- FUNCTION: PLAYS A ROLE IN VALINE AND PYRIMIDINE METABOLISM. BINDS
CC FATTY ACYL-CoA.
CC -1- CATALYTIC ACTIVITY: 2-methyl-3-oxopropanoate + CoA + NAD(+) =
CC propionyl-CoA + CO(2) + NADH.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L08643; AAA30650.1; -;
DR PIR: A46600; A46600.
DR HSSP: P51977; IBXS.
DR InterPro: IPR002086; Aldehyde_dehydr.
DR Pfam: PF00171; aldehyd. 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; FALSE_NEG.
KW Oxidoreductase; NAD; Transit peptide; Mitochondrion.
FT TRANSIT 1 34
FT CHAIN 35 537
FT NP_BIND 263 258
FT ACT_SITE 319 319
FT CONFLICT 126 127
FT SEQUENCE 537 AA; 53062 MW; 69DF39506E2F9C0 CRC64;
Query Match 2.6%; Score 9; DB 1; Length 537;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 13 AVAAAAAL 21
IIIIIIIIII

Db 5 AVAAAAAAL 13

RESULT 5

KPCM_MOUSE STANDARD: PRT: 918 AA.

ID KPCM_MOUSE

AC 062101:

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Protein kinase C, mu type (EC 2.7.1.-) (npkc-mu) (Protein kinase D).

GN PRCM OR PKCM OR PKD.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BALB/C; TISSUE=Lung;

RX MEDLINE=94359973; PubMed=8078925;

RA Valverde A.M., Smet-Smit J., Van Lint J., Rozenburg E.;

RT "Molecular cloning and characterization of protein kinase D: a target for diacylglycerol and phorbol esters with a distinctive catalytic domain.";

RT Proc. Natl. Acad. Sci. U.S.A. 91:8572-8576(1994).

RN [2]

RP PHOSPHORYLATION OF SER-916.

RX MEDLINE=99403106; PubMed=10473617;

RA Matthews S.A., Rozenburg E., Cantrell D.;

RT "Characterization of serine 916 as an in vivo autophosphorylation site for protein kinase D/protein kinase Cmu.";

RT J. Biol. Chem. 274:26543-26549(1999).

CC -1- FUNCTION: THIS IS CALCIUM-INDEPENDENT, PHOSPHOLIPID-DEPENDENT, SERINE- AND THREONINE-SPECIFIC ENZYME.

CC -1- ENZYME REGULATION: ACTIVATED BY DIACYLGLYCEROL AND PHORBOL ESTERS.

CC -1- PTM: Autophosphorylated.

CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG BINDING DOMAINS.

CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC PRC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: Z34524; CAA84283.1; -

DR HSSP: P28867; IPTO.

DR MCD: MGI:99879; prkcm.

DR InterPro: IPR002219; DAG_DE-bind.

DR InterPro: IPR000719; Euk_pkinase.

DR InterPro: IPR001849; PH.

DR InterPro: IPR002290; Ser_thr_pkinase.

DR Pfam: PF00069; pkinase; 1.

DR Pfam: PF00130; DAG_PE-bind; 2.

DR Pfam: PF00169; PH; 1.

DR PRINTS: PR00008; DAGPEDOMAIN.

DR ProDom: PD000001; Euk_pkinase; 1.

DR SMART: SM00109; C1; 2.

DR SMART: SM00233; PH; 1.

DR SMART: SM00220; S_TKC; 1.

DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; 2.

DR PROSITE: PS50081; DAG_PE_BIND_DOM_2; 2.

DR PROSITE: PS50003; PH_DOMAIN; 1.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

DR ATP-binding; Transferase; Serine/threonine-protein kinase;

KW Phorbol-ester binding; zinc; Repeat; Phosphorylation.

FT DOMAIN 145 194 PHORBOL-ESTER AND DAG BINDING 1.

FT DOMAIN 277 326 PHORBOL-ESTER AND DAG BINDING 2.

FT DOMAIN 428 547 PH.

FT DOMAIN 589 845 PROTEIN KINASE.

FT DOMAIN 16 26 POLY-ALA.

FT NP_BIND 198 201 POLY-ARG.

FT BINDING 595 603 ATP (BY SIMILARITY).

FT ACT_SITE 618 618 ATP (BY SIMILARITY).

FT ACT_SITE 712 712 BY SIMILARITY.

FT MOD_RES 916 916 PHOSPHORYLATION (AUTO-).

SO SEQUENCE 918 AA; 102067 MW; 234486180521BD0A CRC64;

Query Match 2.6%; Score 9; DB 1; Length 918;

Best Local Similarity 100.0%; Pred. No. 2.9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13 AVAAAAAAL 21

Db 19 AVAAAAAAL 27

RESULT 6

NI53_HUMAN STANDARD: PRT: 1475 AA.

ID NI53_HUMAN

AC P49790;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Nuclear pore complex protein Nup153 (Nucleoporin Nup153) (153 kDa nucleoporin).

GN NUP153.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94154002; PubMed=8110839;

RA McMorro I., Bastos R., Horton H., Burke B.;

RT "Sequence analysis of a cDNA encoding a human nuclear pore complex protein, hnup153.";

RT Biochim. Biophys. Acta 1217:219-223(1994).

CC -1- FUNCTION: POSSIBLE DNA-BINDING SUBUNIT OF THE NUCLEAR PORE COMPLEX (NPC). THE REPEAT-CONTAINING DOMAIN MAY BE INVOLVED IN ANCHORING COMPONENTS OF THE PORE COMPLEX TO THE PORE MEMBRANE.

CC -1- SUBCELLULAR LOCATION: NUCLEAR PORE COMPLEX. LOCATED TO THE PERINATL RING STRUCTURE OF THE NUCLEOPLASMIC CAGE.

CC -1- DOMAIN: CONTAINS P-X-F-G REPEATS.

CC -1- SIMILARITY: THE REPEAT REGION COMPOSED OF PENTAPEPTIDE REPEATS SEPARATED BY SER/THR-RICH DOMAINS IS SIMILAR TO THAT OF YEAST NUP1, NSP1, POM 121 AND MAMMALIAN P62.

CC -1- SIMILARITY: CONTAINS 4 RANBP2-TYPE ZINC FINGERS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: Z25535; CAA80982.1; -

DR Genew: HGNC:8062; NUP153.

DR MIM: 603948; -

DR InterPro: IPR001876; znF_RanGDP.

DR Pfam: PF00641; znF_RanBP; 4.

DR SMART: SM00547; znF_RB2; 4.

DR PROSITE: PS01358; znF_RANBP2_1; 4.

DR PROSITE: PS50199; znF_RANBP2_2; 4.

KW Nuclear protein; Transport; Repeat; Zinc-finger; DNA-binding.

FT DOMAIN 4 14 GLY-RICH.

FT DOMAIN 443 447 POLY-GLY.

FT ZN_FING 657 687 RANBP2-TYPE 1.

FT 2N_FING 722 751 RANBP2-TYPE 2.
 FT 2N_FING 793 822 RANBP2-TYPE 3.
 FT 2N_FING 851 880 RANBP2-TYPE 4.
 SQ SEQUENCE 1475 AA; 153889 MW; 3CB415A6909DF80E CRC64;

Query Match 2.6%; Score 9; DB 1; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 282 SSVPASSSS 290
 DB 823 SSVPASSSS 831

RESULT 7

ANP3_PSEAM STANDARD; PRT; 37 AA.

AC P02733;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Antifreeze peptide 3.
 OS Pseudopleuronectes americanus (Winter flounder) (Pleuronectes americanus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes; OC Pleuronectoidei; Pleuronectidae; Pseudopleuronectes.
 ON NCB1_TaxId=8265;
 RN [1]

RP MEDLINE=78060969; PubMed=588591;
 RX Devries A.L., Lin Y.;

RT "Structure of a peptide antifreeze and mechanism of adsorption to ice."
 RL Biochem. Biophys. Acta 495:388-392(1977).
 CC -1- FUNCTION: ANTIFREEZE PROTEINS LOWER THE BLOOD FREEZING POINT.
 CC -1- SIMILARITY: BELONGS TO THE TYPE-I AFP FAMILY. TYPE 1 AFP ARE

CC ALANINE-RICH, AMPHIPHILIC AND ALPHAH-HELICAL.
 DR PIR; A03192; FDFL3W.
 DR InterPro; IPR00104; Antifreeze_1.

KW Antifreeze protein; Repeat; Multigene family.
 SQ SEQUENCE 37 AA; 3144 MW; 46AA951A962DECA9 CRC64;

Query Match

Best Local Similarity 100.0%; Score 8; DB 1; Length 37;
 Pred. No. 1.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 AAAAALT 22
 DB 6 AAAAALT 13

RESULT 8

ANP4_PSEAM STANDARD; PRT; 82 AA.

AC P04002;
 DT 23-OCT-1986 (Rel. 02, Created)
 DT 23-OCT-1986 (Rel. 02, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Antifreeze protein A/B precursor.
 OS Pseudopleuronectes americanus (Winter flounder) (Pleuronectes americanus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes; OC Pleuronectoidei; Pleuronectidae; Pseudopleuronectes.
 ON NCB1_TaxId=8265;
 RN [1]

RP SEQUENCE FROM N.A. (PROTEIN A).
 RX MEDLINE=82197490; PubMed=6952188;
 RA Davies P.L., Roach A.H., Hew C.-L.;

RT "DNA sequence coding for an antifreeze protein precursor from winter flounder."
 RT Proc. Natl. Acad. Sci. U.S.A. 79:335-339(1982).

RP SEQUENCE FROM N.A. (PROTEIN A).
 RX MEDLINE=88259236; PubMed=3133486;
 RA Scott G.K., Davies P.L., Kao M.H., Fletcher G.L.;
 RT "Differential amplification of antifreeze protein genes in the pleuronectineae."
 RT J. Mol. Evol. 27:29-35(1988).

RP SEQUENCE FROM N.A. (PROTEIN B).
 RX MEDLINE=84264559; PubMed=6086629;
 RA Davies P.L., Hough C., Scott G.K., Ng N., White B.N., Hew C.-L.;
 RT "Antifreeze protein genes of the winter flounder."
 RT J. Biol. Chem. 259:9241-9247(1984).

RP SEQUENCE FROM N.A.
 RX MEDLINE=92209995; PubMed=1555765;
 RA Davies P.L.;
 RT "Conservation of antifreeze protein-encoding genes in tandem repeats."
 RT Gene 112:163-170(1992).

RP 3D-STRUCTURE MODELING OF 45-81.
 RX MEDLINE=92148833; PubMed=1738160;
 RA Chou K.-C.;
 RT "Energy-optimized structure of antifreeze protein and its binding mechanism."
 RT J. Mol. Biol. 223:509-517(1992).

RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 45-81.
 RX MEDLINE=95281060; PubMed=7760940;
 RA Siebert F., Yang D.S.C.;
 RT "Ice-binding structure and mechanism of an antifreeze protein from winter flounder."
 RT Nature 375:427-431(1995).
 CC -1- FUNCTION: ANTIFREEZE PROTEINS LOWER THE BLOOD FREEZING POINT.
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF PROTEIN A.
 CC -1- SIMILARITY: BELONGS TO THE TYPE-I AFP FAMILY. TYPE 1 AFP ARE

CC ALANINE-RICH, AMPHIPHILIC AND ALPHAH-HELICAL.
 CC THE SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; L00138; AAB55664.1; -
 DR EMBL; L28178; AAB55664.1; JOINED.
 DR EMBL; M62414; AAA45469.1; -
 DR EMBL; X07506; CAA30389.1; -
 DR EMBL; M62416; AAA45471.1; -
 DR EMBL; M62417; AAA45472.1; -
 DR PIR; A03194; FDFLAW.
 DR PIR; A05161; A05161.
 DR PIR; S02326; S02326.
 DR PIR; JS0704; JS0704.
 DR PDB; 1ATF; 15-OCT-94.
 DR PDB; 1WEA; 03-JUN-95.
 DR PDB; 1WFB; 03-JUN-95.
 DR InterPro; IPR00104; Antifreeze_1.
 DR PRINTS; PR00308; ANTIFREEZE1.
 KW Antifreeze protein; Repeat; Multigene family; Signal; 3D-structure.
 FT SIGNAL 1 21
 FT PROPEP 22 44
 FT CHAIN 45 82
 FT VARIANT 36 36
 FT VARIANT 70 70
 FT CONFLICT 24 24
 REMOVED BY A DIPEPTIDYLPEPTIDASE (PROBABLY).
 ANTIFREEZE PROTEIN A/B.
 A -> V.
 A -> D (IN PROTEIN B).
 S -> R (IN REF. 2).

```

FT HELIX 46 80 C2AE7B74C0D46C01 CRC64:
SQ SEQUENCE 82 AA; 7711 MW;

Query Match 2.3%; Score 8; DB 1; Length 82;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AAAAAALT 22
Db 50 AAAAAALT 57

RESULT 9
RS16_RHIL0 STANDARD; PRT; 133 AA.
AC Q98E70;
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 30S ribosomal protein S16.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OX NCBI_TaxID=381;

SEQUENCE FROM N.A.
RC STRAIN=MAF303099;
RC MEDLINE=21082930; PubMed=11214968;
RA Kaneo T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S.,
RA Matsubara A., Iesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsuno M.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
CC -1- SIMILARITY: BELONGS TO THE S16P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP003004; BAB51050.1;
DR InterPro: IPR000307; Ribosomal_S16.
DR Pfam: PF00886; Ribosomal_S16; 1.
DR ProDom: PD003791; Ribosomal_S16; 1.
DR TIGRfams: TIGR00002; S16; 1.
DR PROSITE: PS00732; RIBOSOMAL_S16; FALSE-NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 133 AA; 14358 MW; 63D982AC198E831D CRC64;

Query Match 2.3%; Score 8; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 AAAAAAAA 20
Db 116 AAAAAAAA 123

RESULT 10
RISB_MYCLE STANDARD; PRT; 160 AA.
AC Q9CCP3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.9) (DMRL synthase)

```

```

DE (Lumazine synthase) (Riboflavin synthase beta chain).
GN RIBH OR MLO560.
OS Mycobacterium lepreae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;

SEQUENCE FROM N.A.
RC STRAIN=TN;
RC MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Elgimeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
RA Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrett B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex
CC catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-
CC ribityl-amino-2,4(1H,3H)-pyrimidinone and L-3,4-dihydroxy-2-
CC butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit
CC catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-amino-
CC 2,4(1H,3H)-pyrimidinone with L-3,4-dihydroxy-2-butanone-4-
CC phosphate yielding 6,7-dimethyl-8-lumazine (By similarity).
CC -1- CATALYTIC ACTIVITY: 2 6,7-dimethyl-8-(1-D-ribityl)lumazine +
CC riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.
CC -1- PATHWAY: Riboflavin biosynthesis; last step.
CC -1- SIMILARITY: BELONGS TO THE DMRL SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL583918; CAC30068.1;
DR HSSP: O66529; IHOK.
DR Leproma; MLO560;
DR InterPro: IPR002180; DMRL_synthase.
DR Pfam: PF000885; DMRL_synthase; 1.
DR ProDom: PD003664; DMRL_synthase; 1.
DR TIGRfams: TIGR00114; RfbH; 1.
KW Riboflavin biosynthesis; Transferase; Complete proteome.
FT DOMAIN 140 146 POLY-ALA.
SQ SEQUENCE 160 AA; 16336 MW; EBA48BDEDB8B026C CRC64;

Query Match 2.3%; Score 8; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 AAAAAALT 23
Db 142 AAAAAALT 149

RESULT 11
ZEST_DROMA STANDARD; PRT; 268 AA.
AC Q24597; Q24598; Q24599; Q24600; Q24601; Q24602; Q24603; Q24604;
AC Q27387;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Regulatory protein zeste (Fragment).
GN Z.
OS Drosophila mauritiana (Fruit fly),
OS Drosophila sechellia (Fruit fly), and

```

```

OS Drosophila simulans (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephyroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7238, 7240;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Various strains;
RX MEDLINE=93360802; PubMed=8355601;
RA Hey J., Kilman R.M.;
RT "Population genetics and phylogenetics of DNA sequence variation at
multiple loci within the Drosophila melanogaster species complex.";
RL Mol. Biol. Evol. 10:884-892(1993).
CC -1- FUNCTION: INVOLVED IN TRANSECTATION PHENOMENA (= SYNAPSIS-DEPENDENT
GENE EXPRESSION), WHERE THE SYNAPTIC PAIRING OF CHROMOSOMES
CARRYING GENES WITH WHICH ZESTE INTERACTS INFLUENCES THE
EXPRESSION OF THESE GENES. ZESTE BINDS TO DNA AND STIMULATES
TRANSCRIPTION FROM A NEARBY PROMOTER (BY SIMILARITY).
CC -1- SUBUNIT: SELF-ASSOCIATE FORMING COMPLEXES OF SEVERAL HUNDRED
MONOMERS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L13049; AAA29032.1; -
DR EMBL; L13050; AAA29033.1; -
DR EMBL; L13051; AAA29034.1; -
DR EMBL; L13052; AAA29035.1; -
DR EMBL; L13053; AAA29036.1; -
DR EMBL; L13054; AAA29037.1; -
DR EMBL; L13055; AAA29038.1; -
DR EMBL; L13056; AAA29039.1; -
DR EMBL; L13057; AAA29040.1; -
DR EMBL; L13058; AAA29041.1; -
DR EMBL; L13059; AAA29042.1; -
DR EMBL; L13060; AAA29043.1; -
DR EMBL; L13061; AAA29044.1; -
DR EMBL; L13062; AAA29045.1; -
DR EMBL; L13063; AAA29046.1; -
DR EMBL; L13064; AAA29047.1; -
DR EMBL; L13065; AAA29048.1; -
DR EMBL; L13066; AAA29049.1; -
DR FlyBase; FBgn0012518; Dmdu/z.
DR FlyBase; FBgn0012801; Dsec/z.
DR FlyBase; FBgn0012904; Dsim/z.
KM DNA-binding; Transcription regulation; Nuclear protein.
FT NON_TER 1
FT DNA_BIND 1
FT DOMAIN 97 268
FT FT
FT FT
FT VARIANT 105 106
FT NON_TER 268 268
FT SEQUENCE 268 AA; 30353 MW; F592PF8RA4F2DB47B CRC64;
SO
Query Match 2.3%; Score 8; DB 1; Length 268;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 AVAAAAAA 20
DB 192 AVAAAAAA 199
RESULT 12
ALL4_ASPFU STANDARD; PRT; 286 AA.
ID ALL4_ASPFU

```

```

AC 060024;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Allergen Asp f 4 (Fragment).
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiaceae; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 4202 / AF-102;
RA Hemman S., Nikolazik W.H., Schoeni M.H., Blaser K., Cramer R.;
RT "Diagnosis of allergic bronchopulmonary aspergillosis in patients with
cystic fibrosis by IgE-specific serology with recombinant Aspergillus
fumigatus allergens.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN TRANSECTATION PHENOMENA (= SYNAPSIS-DEPENDENT
GENE EXPRESSION), WHERE THE SYNAPTIC PAIRING OF CHROMOSOMES
CARRYING GENES WITH WHICH ZESTE INTERACTS INFLUENCES THE
EXPRESSION OF THESE GENES. ZESTE BINDS TO DNA AND STIMULATES
TRANSCRIPTION FROM A NEARBY PROMOTER (BY SIMILARITY).
CC -1- SUBUNIT: SELF-ASSOCIATE FORMING COMPLEXES OF SEVERAL HUNDRED
MONOMERS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ001732; CA004959.1; -
KW Allergen.
FT NON_TER 1
FT SEQUENCE 286 AA; 30041 MW; A24D0E1D9AE89758 CRC64;
SO
Query Match 2.3%; Score 8; DB 1; Length 286;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 AVAAAAAA 20
DB 42 AVAAAAAA 49
RESULT 13
CYSD_RHIME
ID CYSD_RHIME STANDARD; PRT; 317 AA.
AC P56892;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sulfate adenylyltransferase subunit 2 (EC 2.7.7.4) (Sulfate adenylylate
transferase) (SAT) (ATP-sulfurylase small subunit).
GN CYSD OR R00943 OR SMC00091.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=93395034; PubMed=10464198;
RA Abola A.P., Willis M.G., Wang R.C., Long S.R.;
RT "Reduction of adenosine-5'-phosphosulfate instead of 3'-
phosphoadenosine-5'-phosphosulfate in cysteine biosynthesis by
Rhizobium meliloti and other members of the family Rhizobiaceae.";
RL J. Bacteriol. 181:5280-5287(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puenhler A., Purnelle B., Ransperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weicher S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont

```

```

RT  Sinorhizobium meliloti strain 1021."
RL  Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC  -1- CATALYTIC ACTIVITY: ATP + sulfate -> diphosphate + adenylylsulfate.
CC  -1- PATHWAY: FIRST STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
CC  REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
CC  BIOSYNTHETIC PATHWAY.
CC  -1- SUBUNIT: HEMERODIMER COMPOSED OF CYSD, THE SMALLER SUBUNIT, AND
CC  CYSDN (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE PAPS REDUCTASE FAMILY. CYSD SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: AF158023; AAC5760.1; -.
DR  EMBL: AL591785; CAC45515.1; -.
DR  InterPro: IPR002500; PAPS_reduct.
DR  Pfam: PF01507; PAPS_reduct. 1.
KM  Cysteine biosynthesis; Transferase; Nucleotidyltransferase;
KW  Complete proteome.
FT  CONFLICT 260
SQ  SEQUENCE 317 AA; 36429 MW; 093F272AD22841CB CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 317;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  243 RDGMILIA 250
Db  234 RDGMILIA 241
|||||||
RESULT 14
YER0_MYCTU STANDARD; PRT; 317 AA.
ID YER0_MYCTU
AC P71761; 053171;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV1480.
GN RV1480 OR MT1527 OR MTV007.27 OR MTCY277.01.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=9825987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feldwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., McLean J., Moule S., Murphy L.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Ormanyan L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and

```

```

RT  laboratory strains."
RL  Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC  -1- SIMILARITY: STRONG. TO M. AVIUM MAJ169.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: AL021184; CA16010.1; -.
DR  EMBL: AE007022; AAK45792.1; ALT_INT.
DR  TIGR: MT1527; -.
DR  Tuberculist: RV1480; -.
DR  InterPro: IPR002881; DUF58.
DR  InterPro: IPR002035; VWF_A.
DR  Pfam: PF01882; DUF58; 1.
DR  SMART: SM00327; VMA; 1.
KM  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 317 AA; 34333 MW; 3D1D060FF5ECCF56 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 317;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  12 LAVAAAAA 19
Db  120 LAVAAAAA 127
|||||||
RESULT 15
CHEB_PSEAE STANDARD; PRT; 368 AA.
ID CHEB_PSEAE
AC 087125;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein-glutamate methyltransferase (EC 3.1.1.61).
GN CHEB OR PA1459.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=99161288; PubMed=10052136;
RA Kato J., Nakamura T., Kuroda A., Ohtake H.;
RT "Cloning and characterization of chemotaxis genes in Pseudomonas
RT aeruginosa."
RL Biosci. Biotechnol. Biochem. 63:155-161(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
RN [3]
RP FUNCTION: INVOLVED IN THE MODULATION OF THE CHEMOTAXIS SYSTEM;
CC CATALYZES THE DEMETHYLATION OF SPECIFIC METHYLGLUTAMATE RESIDUES
CC INTRODUCED INTO THE CHEMORECEPTORS (METHYL-ACCEPTING CHEMOTAXIS
CC PROTEINS) BY CHER (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Protein L-glutamate O4-methyl ester + H(2)O =
CC protein L-glutamate + methanol.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

```

CC -1- DOMAIN: THE N-TERMINAL REGULATORY DOMAIN INHIBITS THE ACTIVITY OF
CC THE C-TERMINAL EFFECTOR DOMAIN.
CC -1- PTH: PHOSPHORYLATED BY CHEA. PHOSPHORYLATION SUPPRESSES THE
CC INHIBITORY ACTIVITY OF THE N-TERMINAL DOMAIN (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RESPONSE REGULATORY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 CHEB-TYPE METHYLESTERASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB012767; BAA33550.1; -;
DR EMBL: AE004573; AAG04848.1; -;
DR HSSP: P04042; 1CHD.
DR InterPro: IPR000673; Cheb_methylast.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF000072; response_reg; 1.
DR ProDom: PD000039; Cheb_methylast; 1.
DR ProDom: PD005328; Cheb_methylast; 1.
DR SMART: SM00448; REC; 1.
DR PROSITE: PS50122; CHEB; 1.
DR PROSITE: PS50110; RESPONSE_REGULATORY; 1.
DR Hydrolyase; Chemotaxis; Sensory transduction; Phosphorylation;
KW Complete proteome.
KW DOMAIN 4 121 RESPONSE REGULATORY.
FT 172 368 CHEB-TYPE METHYLESTERASE.
FT MOD_RES 55 55 PHOSPHORYLATION (BY SIMILARITY).
FT ACT_SITE 192 192 BY SIMILARITY.
FT ACT_SITE 219 219 BY SIMILARITY.
FT ACT_SITE 312 312 BY SIMILARITY.
FT CONFLICT 34 34 G -> A (IN REF. 1).
SQ SEQUENCE 368 AA; 39004 MW; FE4801DC220C613B CRC64;

Query Match 2.3%; Score 8; DB 1; Length 368;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 277 APASSSV 284
|||||||
DB 151 APASSSV 158

Search completed: July 11, 2003, 11:04:40
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:00:00 ; Search time 80 Seconds
(without alignments)
898.881 Million cell updates/sec

Title: US-09-654-652A-3

Sequence: 1 MNKKTKAVKSAALVAAAAA.....AKGAKVNPNGHKRYVNFEEH 349

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_proteic:*
12: sp_virus:*
13: sp_vertibrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	3.2	539	10	Q9SR71
2	10	2.9	250	10	Q9LJ08
3	10	2.9	335	10	Q8RXG3
4	10	2.9	348	10	Q96517
5	10	2.9	349	10	Q80398
6	10	2.9	369	10	Q04322
7	10	2.9	621	2	Q59301
8	9	2.6	87	10	Q9LJ13
9	9	2.6	105	10	Q94G89
10	9	2.6	184	17	Q9HQS5
11	9	2.6	300	2	Q9KWF8
12	9	2.6	321	7	Q9X819
13	9	2.6	337	5	Q95NM9
14	9	2.6	337	5	Q95NM3
15	9	2.6	350	2	Q9KMG4
16	9	2.6	439	17	Q9YCK0

17	9	2.6	518	10	Q9FRC3	Q9FRC3 oryza sativ
18	9	2.6	776	16	Q86516	Q86516 streptomyc
19	9	2.6	340	5	Q9NR03	Q9NR03 aedes aegy
20	9	2.6	1068	5	Q9W3X2	Q9W3X2 drosophila
21	9	2.6	1307	5	Q9V702	Q9V702 drosophila
22	9	2.6	3112	5	Q9NKP1	Q9NKP1 leishmania
23	8	2.3	23	4	Q43519	Q43519 homo sapien
24	8	2.3	37	6	Q8WME7	Q8WME7 sus scrofa
25	8	2.3	38	10	Q9S9D9	Q9S9D9 nicotiana t
26	8	2.3	46	7	P79534	P79534 homo sapien
27	8	2.3	46	7	P79536	P79536 homo sapien
28	8	2.3	46	7	P79538	P79538 homo sapien
29	8	2.3	59	4	Q8TE45	Q8TE45 homo sapien
30	8	2.3	82	12	Q8VAT5	Q8VAT5 white spot
31	8	2.3	82	13	Q9VH51	Q9VH51 pseudopleur
32	8	2.3	82	13	Q99013	Q99013 pseudopleur
33	8	2.3	96	10	Q8W327	Q8W327 oryza sativ
34	8	2.3	100	4	Q9BQR1	Q9BQR1 homo sapien
35	8	2.3	101	10	Q9LGG6	Q9LGG6 oryza sativ
36	8	2.3	101	10	Q9AX16	Q9AX16 oryza sativ
37	8	2.3	105	10	Q943N0	Q943N0 oryza sativ
38	8	2.3	107	4	Q16846	Q16846 homo sapien
39	8	2.3	111	4	Q15587	Q15587 homo sapien
40	8	2.3	134	4	Q43816	Q43816 homo sapien
41	8	2.3	135	4	Q8TD56	Q8TD56 homo sapien
42	8	2.3	137	6	Q29365	Q29365 sus scrofa
43	8	2.3	138	4	Q9NP14	Q9NP14 homo sapien
44	8	2.3	139	6	Q9GLV8	Q9GLV8 equus cabal
45	8	2.3	141	7	Q9GJ06	Q9GJ06 homo sapien

ALIGNMENTS

RESULT 1

Q9SR71 ID Q9SR71 PRELIMINARY: PRT; 539 AA.
AC Q9SR71; 01-MAY-2000 (TREMURel. 13, Created)
DT 01-MAY-2000 (TREMURel. 13, Last sequence update)
DT 01-MAY-2002 (TREMURel. 20, Last annotation update)
DE T22K18.10 protein (Putative TBP-associated 58 kDa subunit
DE protein).
GN T22K18.10 OR TAFII158.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA Rensing C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC T22K18 genomic sequence.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC "Arabidopsis thaliana TAFII158, a member of the TFIID complex.";
RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AC010927; AF0417.1; -;
DR EMBL: AF344878; AK28289.1; -;
DR InterPro: IPR004822; Histone-core.
DR InterPro: IPR002965; P-rich_extensn.
DR PRINTS: PR01217; PRICHEXTENS.
DR PRODOM: PD012998; TFIID_sub.1.
SQ SEQUENCE 539 AA; 57710 MW; 72597A2E1B57EE9B CRC64;
Query Match 3.2%; Score 11; DB 10; Length 539;
Best local Similarity 100.0%; Pred. No. 0.081;

Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 273 SSSSAPASSS 283
 Db 117 SSSSAPASSS 127

RESULT 2

09LJ08 PRELIMINARY; PRT; 250 AA.
 AC 09LJ08;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ESTS AU067919(C10906).
 OS Oryza sativa (rice).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: Liliopsida; Poales; Poaceae;
 OC Eriarthroidae; Oryzaceae; Oryza.
 NCBI_TaxID=4530;
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
 clone: P0667A10.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP001073; BAB89582.1; -.
 DR ANU-2DPAGE: 09LJ08; -.
 DR InterPro: IPR000010; Cystatin.
 DR InterPro: IPR003243; Cystatin_C/M.
 DR Pfam: PF00031; Cystatin_2.
 DR ProDom: PD001231; Cystatin_C/M; 1.
 DR SMART: SM00043; CY; 1.
 DR PROSITE: PS00287; CYPSTATIN; UNKNOWN_1.
 SQ SEQUENCE 250 AA: 27269 MW: 6F2A/DHE8BA66F CRC64;

Query Match 2.9%; Score 10; DB 10; Length 250;
 Best Local Similarity 100.0%; Pred. No. 0.38;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LAVAAAAAAL 21
 Db 24 LAVAAAAAAL 33

RESULT 3

08RXG3 PRELIMINARY; PRT; 335 AA.
 AC 08RXG3;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE MAP kinase kinase 5.
 GN A7G621220.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota: Viridiplantae: Streptophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carlini P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kanliya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shin P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY081272; AAL91161.1; -.
 KW kinase

SO SEQUENCE 335 AA: 37048 MW: 0D65B7F67D698521 CRC64;

Query Match 2.9%; Score 10; DB 10; Length 335;
 Best Local Similarity 100.0%; Pred. No. 0.49;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 SSSSAPASSS 282
 Db 33 SSSSAPASSS 42

RESULT 4

096517 PRELIMINARY; PRT; 348 AA.
 AC 096517;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE MAP kinase kinase alpha protein kinase (EC 2.7.1.37).
 GN MAP2KALPHA.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97092864; Pubmed=8938426;
 RA Jouanin S., Hamal A., Kreis M., Henry Y.;
 RT "Molecular cloning of the asparagine synthetase gene from Asparagus
 officinalis L.";
 RL Plant Physiol. 112:1397-1397(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hamal A., Jouanin S., Leprince S., Kreis M., Henry Y.;
 RT "Molecular characterization and expression of an Arabidopsis Thaliana
 L. MAP kinase kinase cDNA AtMAP2Kalpha.";
 RL Plant Sci. 140:41-52(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Jouanin S., Leprince A.S., Hamal A., Kreis M., Henry Y.;
 RT "Plant MAP kinase signalling pathways in the limekiln.";
 RL Adv. Bot. Res. 30:0-0(2000).
 DR EMBL: Y07694; CAA68958.1; -.
 DR HSSP: P24941; 1B0H.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase_1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Kinase; Transferase.

SO SEQUENCE 348 AA: 38329 MW: C18C987C6B7AF5F5 CRC64;

Query Match 2.9%; Score 10; DB 10; Length 348;
 Best Local Similarity 100.0%; Pred. No. 0.5;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 SSSSAPASSS 282
 Db 46 SSSSAPASSS 55

RESULT 5

080398 PRELIMINARY; PRT; 348 AA.
 AC 080398;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE MAP kinase kinase 5.
 GN MKK5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:

SO SEQUENCE 348 AA: 37048 MW: 0D65B7F67D698521 CRC64;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Shinozaki K., Ichimura K., Mizoguchi T.;
RT "Molecular cloning and characterization of cDNAs encoding mitogen-
RT activated protein kinase kinase in Arabidopsis thaliana.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Nakamura Y.;
RX MEDLINE=20277480; PubMed=10819329;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT clones.";
RT features of the regions of 4,504,864 bp covered by sixty pl and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AB015316; BAA2831.1; -;
DR EMBL: AB023045; BAB01714.1; -;
DR HSSP: P24941; 1A01.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 348 AA; 38329 MW; 216C96728F34B55 CRC64;

Query Match 2.9%; Score 10; DB 10; Length 348;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 273 SSSAPASS 282
Db 46 SSSAPASS 55
|||||
PRELIMINARY; PRT; 369 AA.
ID 004322
AC 004322;
DT 01-JUL-1997 (TREMUREL. 04, Created)
DT 01-JUL-1997 (TREMUREL. 04, Last sequence update)
DT 01-MAR-2002 (TREMUREL. 20, Last annotation update)
DE MYB-related transcription activator (MYBST1) ISOLOG (MYBST1).
GN T02004.16.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Rounsley S.D., Lin X., Ketchum K.A., Phillips C.A., Brandon R.C.,
RA Fuhrmann J.E., White O., Kerlavage A.R., Adams M.D., Somerville C.R.,
RA Venter J.C.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA MEDLINE=20277480; PubMed=10819329;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT clones.";
RT features of the regions of 4,504,864 bp covered by sixty pl and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL: AC001645; AAE3650.1; -;
DR EMBL: AB023046; BAE01274.1; -;
DR InterPro: IPR001005; Myb_DNA_binding.
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam: PF00249; myb_DNA-binding; 1.
DR SMART: SM00395; SANT; 1.
DR SMART: SM00343; Znf_C2HC; 1.
DR PROSITE: PS00900; MYB_3; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 369 AA; 40043 MW; DCC039507044AC9E CRC64;

Query Match 2.9%; Score 10; DB 10; Length 369;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 SALAVAAA 19
Db 52 SALAVAAA 61
|||||
PRELIMINARY; PRT; 621 AA.
ID 059301
AC 059301;
DT 01-NOV-1996 (TREMUREL. 01, Created)
DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)
DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
DE Endo-beta-1,4-xylanase precursor (Ec 3.2.1.8).
OS Cellvibrrio mixtus.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Cellvibrrio.
OX NCBI_TaxID=39650;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MIXTUS;
RX MEDLINE=96077124; PubMed=7492333;
RA Millward-Sadler S.J., Davidson K., Hazlewood G.P., Black G.W.,
RA Gilbert H.J., Clarke J.H.;
RT "Novel cellulose-binding domains, NodB homologues and conserved
RT modular architecture in xylanases from the aerobic soil bacteria
RT Pseudomonas fluorescens subsp. cellulosa and Cellvibrrio mixtus.";
RL Biochem. J. 312:39-48(1995).
DR EMBL: Z48926; CAA88762.1; -;
DR HSSP: P14768; ICIX.
DR InterPro: IPR005086; CBM_15.
DR InterPro: IPR001000; Glyco_hydro_10.
DR Pfam: PF03426; CBM_15; 1.
DR Pfam: PF00331; Glyco_hydro_10; 1.
DR PRINTS: PR00134; GLYHDLASE10.
DR PROSITE: PS00591; GLYCOSYL_HYDROL_F10; UNKNOWN_1.
KW Glycosidase; Hydrolase; Signal; Xylan degradation.
FT SIGNAL 19
FT SIGNAL 19
SQ SEQUENCE 621 AA; 64929 MW; 830E5B95D2C3A88 CRC64;

Query Match 2.9%; Score 10; DB 2; Length 621;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 274 SSSAPASS 283
Db 89 SSSAPASS 98
|||||

RESULT 8

09L113 ID 09L113 PRELIMINARY: PRT: 87 AA.
 AC 09L113;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE Hypothetical protein (OSJNB0086P08.3 protein).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone: p0708G02.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone: OSJNB0086P08.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP001539; BAA92926.1;
 DR EMBL: AP002855; BAB7194.1;
 DR EMBL: AF002855; BAB7194.1;
 SO SEQUENCE 87 AA; 9123 MW; F3ACD82BA63824EE CRC64;

Query Match 2.6%; Score 9; DB 10; Length 87;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 LAVAAAAA 20
 ||||||||
 DB 22 LAVAAAAA 30

RESULT 9

094GH9 ID 094GH9 PRELIMINARY: PRT: 105 AA.
 AC 094GH9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 10.1 kDa protein.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, NIPPONBARE;
 RA Buell C.R., Yuan O., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
 RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsirlin T.,
 RA Riggs F., Hsiao J., Zismann V., Blunt S., Pal G., VanAken S.E.,
 RA Utecher O.T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
 RA White O., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNB0093E13 genomic sequence.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC091123; AAK72898.1;
 DR EMBL: AC091123; AAK72898.1;
 SO SEQUENCE 105 AA; 10143 MW; 8EF84EE04A339F7 CRC64;

Query Match 2.6%; Score 9; DB 10; Length 105;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 LAVAAAAA 20
 ||||||||
 DB 11 LAVAAAAA 19

RESULT 10

09H0S5 ID 09H0S5 PRELIMINARY: PRT: 184 AA.
 AC 09H0S5;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE Vng1026h.
 DE Vng1026h.
 GN VNG1026H.
 OS Halobacterium sp. (strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Weller R., Goo Y.A.,
 RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbacher T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omar A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RT "Genome sequence of Halobacterium species NRC-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 DR EMBL: AE005037; AAG19438.1;
 DR EMBL: AE005037; AAG19438.1;
 KW Complete proteome.
 SO SEQUENCE 184 AA; 18956 MW; B120BC2158DD9095 CRC64;

Query Match 2.6%; Score 9; DB 17; Length 184;
 Best Local Similarity 100.0%; Pred. No. 2.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 LAVAAAAA 20
 ||||||||
 DB 137 LAVAAAAA 145

RESULT 11

09KMF8 ID 09KMF8 PRELIMINARY: PRT: 300 AA.
 AC 09KMF8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Mevalonate diphosphate decarboxylase.
 GN MDPD.
 OS Kitasatospora grisicola (Streptomyces griseoliosporus).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Kitasatospora.
 OX NCBI_TaxID=2064;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dairi T.;
 RT "Cloning of a gene cluster encoding enzymes responsible for the
 mevalonate pathway from a terpenoid antibiotic-producing Streptomyces
 strain.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB037907; BAB07618.1;
 DR InterPro: IPR001745; GHMPkinse_ATP.
 DR Pfam: PF00288; GHMP_kinases; 1.
 DR TIGRFAMs: TIGR01240; mevdpcdecap; 1.
 SO SEQUENCE 300 AA; 30865 MW; 17D9E9D07781997 CRC64;

Query Match 2.6%; Score 9; DB 2; Length 300;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ALAVAAAA 19
|11111111|
Db 78 ALAVAAAA 86

RESULT 12

Q9XS19 PRELIMINARY; PRT; 321 AA.

AC Q9XS19; (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE MHC class I related protein (Fragment).
GN MIC1.

OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Papio.
OX NCBI_Taxid=9557;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=2093998; PubMed=10630302;
RA Seo J.W., Bontrop R., Walter L., Guenther E.;
RT "Major histocompatibility complex-linked MIC genes in rhesus macaques and other primates."
RT Immunogenetics 50:358-362(1999).
RL -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-MICROGLOBULIN) (BY SIMILARITY).
CC EMBL: AJ242442; CAB46518.1; -;
DR InterPro: IPR003597; Iq-cl.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR001039; MHC_I.
DR Pfam: PF00047; Iq; 1.
DR Pfam: PF00129; MHC_I; 1.
DR PRINTS: PR01638; MHCCLASSI.
DR PRODOM: PD000050; MHC_I; 1.
DR SMART: SM00407; IGcl; 1.
KW Glycoprotein; Transmembrane.

FT NON_TER 1
FT 321 321
SQ SEQUENCE 321 AA; 36039 MW; 8B2579752CE6B65F CRC64;
Query Match 2.6%; Score 9; DB 7; Length 321;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LAVAAAAA 20
|11111111|
Db 291 LAVAAAAA 299

RESULT 13

Q95NU9 PRELIMINARY; PRT; 337 AA.

AC Q95NU9; (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hairy.
GN H OR CG6494.

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-R3-24, R3-2, R3-19, R3-53, R3-48, R3-95, AND R3-74;
RA Robin C., Lyman R.F., Long A.D., Langley C.H., Mackay T.F.C.;
RT "Polymorphism at hairy associated with variation in Drosophila sensory

RT bristle number."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY055842; AL117776.1; -;
DR EMBL: AY055836; AL117770.1; -;
DR EMBL: AY055837; AL117771.1; -;
DR EMBL: AY055838; AL117772.1; -;
DR EMBL: AY055839; AL117773.1; -;
DR EMBL: AY055840; AL117774.1; -;
DR EMBL: AY055841; AL117775.1; -;
DR FlyBase: FBgn0001168; h.
DR InterPro: IPR001092; HLH_basic.

DR Pfam: PF00010; HLH; 1.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 337 AA; 37005 MW; 49BECAPFE2D69FC4 CRC64;
Query Match 2.6%; Score 9; DB 5; Length 337;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LAVAAAAA 20
|11111111|
Db 238 LAVAAAAA 246

RESULT 14

Q95NH3 PRELIMINARY; PRT; 337 AA.

AC Q95NH3; (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hairy.
GN H OR CG6494.

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-R3-6, R3-105, AND R3-107;
RA Robin C., Lyman R.F., Long A.D., Langley C.H., Mackay T.F.C.;
RT "Polymorphism at hairy associated with variation in Drosophila sensory bristle number."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY055835; AL117769.1; -;
DR EMBL: AY055833; AL117767.1; -;
DR EMBL: AY055834; AL117768.1; -;
DR FlyBase: FBgn0001168; h.
DR InterPro: IPR001092; HLH_basic.

DR Pfam: PF00010; HLH; 1.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 337 AA; 37021 MW; E08B90F942B245C0 CRC64;

Query Match 2.6%; Score 9; DB 5; Length 337;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LAVAAAAA 20
|11111111|
Db 238 LAVAAAAA 246

RESULT 15

Q9KMG4 PRELIMINARY; PRT; 350 AA.

AC Q9KMG4; (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Diphosphomevalonate decarboxylase.
OS Streptomyces sp. (strain CL190).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomyclineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=933372;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CL190;
 RX MEDLINE=20353446; PubMed=10894721;
 RA Takagi M., Kuzuyama T., Takahashi S., Seto H.;
 RT "A gene cluster for the mevalonate pathway from Streptomyces sp.
 strain CL190.";
 RL J. Bacteriol. 182:4153-4157(2000).
 DR EMBL; AB037666; BAB0791.1; -;
 DR InterPro; IPR001745; GHMPkinse_ATP.
 DR InterPro; IPR001459; Mew_gal_kin.
 DR Pfam; PF00288; GHMP_kinases; 1.
 DR PRINTS; PRO0959; MEYGALKINASE.
 DR TIGRfams; TIGR01240; mewpdecarb; 1.
 SO SEQUENCE 350 AA; 36409 MW; 1CE61F306ACDC7CF CRC64;

Query Match 2.68; Score 9; DB 2; Length 350;

Best Local Similarity 100.0%; Pred. No. 4.6;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ALAVAAAAA 19
 |||||
 Db 130 ALAVAAAAA 138

Search completed: July 11, 2003, 11:06:08
 Job time : 82 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:02:55 ; Search time 26 Seconds
(without alignments)
394.946 Million cell updates/sec

Title: US-09-654-652A-3

Perfect score: 349

Sequence: 1 MNIKRTAVKSAALAVAAAAA.....AKGAKVNPNGHKRYRNFEEH 349

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	228	65.3	228	3	US-09-286-690-12
2	8	2.3	9	3	US-08-159-339A-152
3	8	2.3	37	2	US-08-180-524-1
4	8	2.3	37	2	US-08-180-524-8
5	8	2.3	37	2	US-08-180-524-9
6	8	2.3	37	2	US-08-975-166-1
7	8	2.3	37	2	US-08-975-166-8
8	8	2.3	37	2	US-08-975-166-9
9	8	2.3	37	4	US-09-117-121-39
10	8	2.3	37	4	US-09-117-121-42
11	8	2.3	38	2	US-09-344-529-4
12	8	2.3	38	2	US-07-814-420-1
13	8	2.3	38	2	US-07-814-421-1
14	8	2.3	340	5	PCT-US96-02331-13
15	8	2.3	340	5	US-09-100-664A-2
16	8	2.3	340	3	US-09-100-664A-3
17	8	2.3	440	3	US-09-100-664A-4
18	8	2.3	440	4	US-09-335-983-2
19	8	2.3	440	4	US-09-335-983-3
20	8	2.3	440	4	US-09-335-983-4
21	8	2.3	461	1	US-08-194-338-4
22	8	2.3	461	1	US-07-881-075-1
23	8	2.3	485	1	US-08-120-827-1
24	8	2.3	485	1	US-08-178-675-1
25	8	2.3	497	1	US-08-009-675-4
26	8	2.3	556	4	US-09-795-691-2
27	8	2.3	655	4	US-08-857-076-57

28	8	2.3	738	3	US-08-864-038A-3	Sequence 3, Appl
29	8	2.3	760	1	US-08-195-152-2	Sequence 2, Appl
30	8	2.3	342	5	PCT-US96-02331-15	Sequence 15, Appl
31	8	2.3	3969	4	US-08-061-576-5	Sequence 5, Appl
32	7	2.0	14	2	US-08-503-226B-42	Sequence 42, Appl
33	7	2.0	14	3	US-08-721-458B-42	Sequence 42, Appl
34	7	2.0	20	6	5422425-10	Patent No. 5422425
35	7	2.0	22	3	US-08-516-859A-103	Sequence 103, App
36	7	2.0	22	4	US-09-586-472-103	Sequence 103, App
37	7	2.0	22	4	US-09-528-706-103	Sequence 103, App
38	7	2.0	30	6	5422425-9	Patent No. 5422425
39	7	2.0	34	1	US-08-425-069-52	Sequence 52, Appl
40	7	2.0	34	2	US-08-317-844B-52	Sequence 52, Appl
41	7	2.0	92	4	US-09-344-529-2	Sequence 2, Appl
42	7	2.0	109	4	US-08-325-832A-144	Sequence 144, App
43	7	2.0	130	2	US-08-853-659A-39	Sequence 39, Appl
44	7	2.0	171	2	US-08-609-049A-22	Sequence 22, Appl
45	7	2.0	171	4	US-09-170-996-22	Sequence 22, Appl

ALIGNMENTS

```
RESULT 1
US-09-286-690-12
; Sequence 12, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; EARLIER FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Fibrobacter succinogenes
US-09-286-690-12

Query Match      65.3%: Score 228; DB 3; Length 228;
Best Local Similarity 100.0%: Pred. No. 4.8e-215;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MNIKRTAVKSAALAVAAAAAALTTNVSADKSGAELTYLLEVOYQKFEARKKMAAGTGS 60
        |||||||
Db       1 MNIKRTAVKSAALAVAAAAAALTTNVSADKSGAELTYLLEVOYQKFEARKKMAAGTGS 60

QY      61 SMFLYQNGSILAGRPVEVDIEYLGNKPSFOSNITTGAGAKTSEKHHAVSPAADQA 120
        |||||||
Db       61 SMFLYQNGSILAGRPVEVDIEYLGNKPSFOSNITTGAGAKTSEKHHAVSPAADQA 120

QY      121 FHRYGLEMTNRYRWYVDGQEVKRTBEGQVSNLTGTGGLFNNLMSSESAWVQFDESKL 180
        |||||||
Db       121 FHRYGLEMTNRYRWYVDGQEVKRTBEGQVSNLTGTGGLFNNLMSSESAWVQFDESKL 180

QY      181 PLFOFINWVKYVYTPQGGSGSDFTLDWTNDNFTDGRGSRMGKGDWTF 228
        |||||||
Db       181 PLFOFINWVKYVYTPQGGSGSDFTLDWTNDNFTDGRGSRMGKGDWTF 228

RESULT 2
US-08-159-339A-152
; Sequence 152, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
```

APPLICANT: Kubo, Ralph T.
APPLICANT: Grey, Howard M.
APPLICANT: Sette, Alessandro
APPLICANT: Celis, Esteban
TITLE OF INVENTION: HLA Binding peptides and Their
NUMBER OF SEQUENCES: 1254
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/159,339A
FILING DATE: 29-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,666
FILING DATE: 07-AUG-1992
APPLICATION NUMBER: US 08/027,746
FILING DATE: 05-MAR-1993
APPLICATION NUMBER: US 08/103,396
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 018623-005030US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX:
INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-159-339A-152

Query Match 2.3%; Score 8; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 AVAAAAA 20
|||||||
DB 1 AVAAAAA 8

RESULT 3
US-08-180-524-1
Sequence 1, Application US/08180524
Patent No. 5849537
GENERAL INFORMATION:
APPLICANT: Tripp, Matthew
APPLICANT: Lusk, Lance
APPLICANT: Rhodes, Thomas
APPLICANT: Hulge, Nick
APPLICANT: Chicoe, Etzer
APPLICANT: Barney, Michael C.
APPLICANT: Bower, Patricia A.
APPLICANT: Cronan, Charles L.
TITLE OF INVENTION: METHOD OF EXPRESSING ANTIFREEZE PROTEINS
NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
ADDRESSEE: Thad Kryshak, Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: USA
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MSWORD Version 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/180,524
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/917,216
FILING DATE:
APPLICATION NUMBER: US 07/486,333
FILING DATE: 28-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/409,217
FILING DATE: 19-SEP-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kryshak, Thad
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 66-005-9234-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5707
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudopleuronectes americanus
US-08-180-524-1

Query Match 2.3%; Score 8; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AAAAALT 22
|||||||
DB 6 AAAAALT 13

RESULT 4
US-08-180-524-8
Sequence 8, Application US/08180524
Patent No. 5849537
GENERAL INFORMATION:
APPLICANT: Tripp, Matthew
APPLICANT: Lusk, Lance
APPLICANT: Rhodes, Thomas
APPLICANT: Hulge, Nick
APPLICANT: Chicoe, Etzer
APPLICANT: Barney, Michael C.
APPLICANT: Bower, Patricia A.
APPLICANT: Cronan, Charles L.
TITLE OF INVENTION: METHOD OF EXPRESSING ANTIFREEZE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thad Kryshak, Quarles & Brady
STREET: 411 East Wisconsin Avenue

CITY: Milwaukee
STATE: Wisconsin
COUNTRY: USA
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MSWORD Version 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/180,524
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/917,216
FILING DATE:
APPLICATION NUMBER: US/07/486,333
FILING DATE: 28-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/409,217
FILING DATE: 19-SEP-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kryshak, Thad
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 66-005-9234-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5707
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-180-524-8

Query Match 2.3%; Score 8; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AAAAAALT 22
|||||||
DB 6 AAAAAALT 13

RESULT 5
US-08-180-524-9
Sequence 9, Application US/08180524
Patent No. 5849537
GENERAL INFORMATION:
APPLICANT: Tripp, Matthew
APPLICANT: Lusk, Lance
APPLICANT: Rhodes, Thomas
APPLICANT: Hulge, Nick
APPLICANT: Kot, Edward
APPLICANT: Chicoye, Etzer
APPLICANT: Barney, Michael C.
APPLICANT: Bower, Patricia A.
APPLICANT: Cronan, Charles L.
TITLE OF INVENTION: METHOD OF EXPRESSING ANTIFREEZE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thad Kryshak, Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: USA
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MSWORD Version 5.0
CURRENT APPLICATION DATA:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MSWORD Version 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/180,524
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/917,216
FILING DATE:
APPLICATION NUMBER: US/07/486,333
FILING DATE: 28-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/409,217
FILING DATE: 19-SEP-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kryshak, Thad
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 66-005-9234-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5707
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-180-524-9

Query Match 2.3%; Score 8; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AAAAAALT 22
|||||||
DB 6 AAAAAALT 13

RESULT 6
US-08-975-166-1
Sequence 1, Application US/08975166
Patent No. 5928877
GENERAL INFORMATION:
APPLICANT: Tripp, Matthew
APPLICANT: Lusk, Lance
APPLICANT: Rhodes, Thomas
APPLICANT: Hulge, Nick
APPLICANT: Kot, Edward
APPLICANT: Chicoye, Etzer
APPLICANT: Barney, Michael C.
APPLICANT: Bower, Patricia A.
APPLICANT: Cronan, Charles L.
TITLE OF INVENTION: METHOD OF EXPRESSING ANTIFREEZE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thad Kryshak, Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: USA
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MSWORD Version 5.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/975,166
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/180,524
FILING DATE:
APPLICATION NUMBER: US/07/917,216
FILING DATE:
APPLICATION NUMBER: US/07/486,333
FILING DATE: 28-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/409,217
FILING DATE: 19-SEP-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kryshak, Thad
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 66-005-9234-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5707
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudopleuronectes americanus
US-08-975-166-1

Query Match 2.3% Score 8; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AAAAALT 22
|||||||
DB 6 AAAAALT 13

RESULT 7
US-08-975-166-8
Sequence 8, Application US/08975166
Patent No. 5928877
GENERAL INFORMATION:
APPLICANT: Tripp, Matthew
APPLICANT: Lusk, Lance
APPLICANT: Rhodes, Thomas
APPLICANT: Hulge, Nick
APPLICANT: Kot, Edward
APPLICANT: Chicoye, Etzer
APPLICANT: Barney, Michael C.
APPLICANT: Bower, Patricia A.
APPLICANT: Cronan, Charles L.
TITLE OF INVENTION: METHOD OF EXPRESSING ANTIFREEZE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thad Kryshak, Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: USA
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MSWORD Version 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,166

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/180,524
FILING DATE:
APPLICATION NUMBER: US/07/917,216
FILING DATE:
APPLICATION NUMBER: US 07/486,333
FILING DATE: 28-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/409,217
FILING DATE: 19-SEP-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kryshak, Thad
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 66-005-9234-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5707
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-975-166-8

Query Match 2.3% Score 8; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AAAAALT 22
|||||||
DB 6 AAAAALT 13

RESULT 8
US-08-975-166-9
Sequence 9, Application US/08975166
Patent No. 5928877
GENERAL INFORMATION:
APPLICANT: Tripp, Matthew
APPLICANT: Lusk, Lance
APPLICANT: Rhodes, Thomas
APPLICANT: Hulge, Nick
APPLICANT: Kot, Edward
APPLICANT: Chicoye, Etzer
APPLICANT: Barney, Michael C.
APPLICANT: Bower, Patricia A.
APPLICANT: Cronan, Charles L.
TITLE OF INVENTION: METHOD OF EXPRESSING ANTIFREEZE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thad Kryshak, Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: USA
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MSWORD Version 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,166
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/180,524
FILING DATE: US/07/917,216
APPLICATION NUMBER: US/07/917,216
FILING DATE: US/07/917,216
APPLICATION NUMBER: US/07/486,333
FILING DATE: 28-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/409,217
FILING DATE: 19-SEP-1989
ATTORNEY/AGENT INFORMATION:
NAME: Krystak, Thad
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 66-005-9234-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5707
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-975-166-9

Query Match 2.3%; Score 8; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AAAAAALT 22
|||||||
DB 6 AAAAAALT 13

RESULT 9
US-09-117-121-39
Sequence 39, Application US/09117121
Patent No. 6307020
GENERAL INFORMATION:
APPLICANT: Hew, Choy
APPLICANT: Gong, Zhiluan
TITLE OF INVENTION: Intracellular Antifreeze Polypeptides
TITLE OF INVENTION: and Nucleic Acids
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/117,121
FILING DATE: 20-NOV-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CA97/00062
FILING DATE: 30-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 016252-001610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: -
LOCATION: 1..37
OTHER INFORMATION: /note="HPLC-6"
US-09-117-121-39

Query Match 2.3%; Score 8; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AAAAAALT 22
|||||||
DB 6 AAAAAALT 13

RESULT 10
US-09-117-121-42
Sequence 42, Application US/09117121
Patent No. 6307020
GENERAL INFORMATION:
APPLICANT: Hew, Choy
APPLICANT: Gong, Zhiluan
TITLE OF INVENTION: Intracellular Antifreeze Polypeptides
TITLE OF INVENTION: and Nucleic Acids
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/117,121
FILING DATE: 20-NOV-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CA97/00062
FILING DATE: 30-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 016252-001610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-117-121-42

Query Match 2.3%; Score 8; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AAAAAALT 22
|||||||
DB 6 AAAAAALT 13

RESULT 11
US-09-344-529-4
Sequence 4, Application US/09344529
Patent No. 6429293
GENERAL INFORMATION:
APPLICANT: Hew, Choy L.
TITLE OF INVENTION: HSC Research and Development Limited Partnership
FILE REFERENCE: 016252-00262005
CURRENT APPLICATION NUMBER: US/09/344,529
CURRENT FILING DATE: 1999-06-24
EARLIER APPLICATION NUMBER: US 60/090,794
EARLIER FILING DATE: 1998-06-26
EARLIER APPLICATION NUMBER: US 60/095,713
EARLIER FILING DATE: 1998-08-07
NUMBER OF SEQ ID NOS: 19
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 37
TYPE: PRT
ORGANISM: Pleuronectes americanus
FEATURE:
OTHER INFORMATION: Winter flounder liver-type antifreeze polypeptide
US-09-344-529-4
Query Match 2.3%; Score 8; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 AAAAAALT 22
| | | | | | | |
DB 6 AAAAAALT 13
RESULT 12
US-07-814-220-1
Sequence 1, Application US/07814220
Patent No. 5925540
GENERAL INFORMATION:
APPLICANT: Caceci, Thomas
APPLICANT: Toth, Thomas E.
APPLICANT: Szumanski, Maria B.W.
TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND
TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: WHITHAM, CURTIS & WHITHAM
STREET: Reston Inl. Center, 11800 Sunrise Valley Dr.,
STREET: Suite 900
CITY: Reston
STATE: VA
COUNTRY: USA
ZIP: 20191
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/814,220
FILING DATE: 23-DEC-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/588,437
FILING DATE: 25-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Whitham, Michael E.
REGISTRATION NUMBER: 32,635
REFERENCE/DOCKET NUMBER: CIT. 016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-391-2510

TELEFAX: 703-391-9035
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-07-814-220-1
Query Match 2.3%; Score 8; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 AAAAAALT 22
| | | | | | | |
DB 6 AAAAAALT 13
RESULT 13
US-07-812-421-1
Sequence 1, Application US/07812421
Patent No. 5932697
GENERAL INFORMATION:
APPLICANT: Caceci, Thomas
APPLICANT: Toth, Thomas E.
APPLICANT: Szumanski, Maria B.W.
TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND
TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: WHITHAM, CURTIS & WHITHAM
STREET: Reston Inl. Center, 11800 Sunrise Valley Dr.,
STREET: Suite 900
CITY: Reston
STATE: VA
COUNTRY: USA
ZIP: 20191
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/812,421
FILING DATE: 23-DEC-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/588,437
FILING DATE: 25-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Whitham, Michael E.
REGISTRATION NUMBER: 32,635
REFERENCE/DOCKET NUMBER: CIT. 016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-391-2510
TELEFAX: 703-391-9035
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-07-812-421-1
Query Match 2.3%; Score 8; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 AAAAAALT 22
| | | | | | | |
DB 6 AAAAAALT 13

RESULT 14
PCT-US96-02331-13
Sequence 13, Application PC/TUS9602331
GENERAL INFORMATION:
APPLICANT: The Board of Trustees of the Leland Stanford Junior
APPLICANT: University
APPLICANT: Board of Regents, The University of Texas System
TITLE OF INVENTION: Methods and Compositions for Altering
TITLE OF INVENTION: Sexual Behavior
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/02331
FILING DATE: 09-FEB-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,495
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8600-0153.41
TELEPHONE: (415) 324-0880
TELECOMMUNICATION INFORMATION:
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-02331-13

Query Match 2.3%; Score 8; DB 5; Length 340;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 AVAAAAA 20
|||||||
Db 246 AVAAAAA 253

RESULT 15
US-09-100-664A-2
Sequence 2, Application US/09100664A
Patent No. 6057129
GENERAL INFORMATION:
APPLICANT: YOUNG, MICHAEL W.
APPLICANT: KLOSS, BRIAN
APPLICANT: BLAU, JUSTIN
APPLICANT: PRICE, JEFFREY
TITLE OF INVENTION: A NOVEL CLOCK GENE AND METHODS OF USE
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA

ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,664A
FILING DATE: 19-JUN-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-221
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-100-664A-2

Query Match 2.3%; Score 8; DB 3; Length 440;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 AVAAAAA 20
|||||||
Db 325 AVAAAAA 332

Search completed: July 11, 2003, 11:07:29
Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:06:15 ; Search time 1523 Seconds
(Without alignments)
26.465 Million cell updates/sec

Title: US-09-654-652A-3
Perfect score: 349
Sequence: 1 MNIKKTAVKSAALAVAAAAA.....AKGAKVNPNGHKRYRVNEH 349

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 440803 seqs, 115490842 residues

Word size : 0

Total number of hits satisfying chosen parameters: 440803

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	2.9	241	9 US-10-188-947-4	Sequence 4, Appli
2	10	2.9	369	9 US-10-278-173-156	Sequence 156, App
3	8	2.3	20	10 US-09-864-761-38908	Sequence 38908, A
4	8	2.3	149	10 US-09-864-761-43779	Sequence 43779, A
5	8	2.3	152	9 US-10-153-668-250	Sequence 250, App
6	8	2.3	287	9 US-10-156-761-9684	Sequence 9684, App
7	8	2.3	320	9 US-10-156-761-13308	Sequence 13308, A
8	8	2.3	338	9 US-10-086-156-6	Sequence 6, Appli
9	8	2.3	338	9 US-10-071-458-4	Sequence 4, Appli
10	8	2.3	374	9 US-10-286-264-112	Sequence 112, App
11	8	2.3	374	9 US-09-934-455-304	Sequence 304, App
12	8	2.3	427	9 US-10-156-761-12190	Sequence 12190, A
13	8	2.3	440	9 US-10-108-605-199	Sequence 199, App
14	8	2.3	440	9 US-10-108-605-219	Sequence 219, App
15	8	2.3	440	9 US-10-108-605-243	Sequence 243, App
16	8	2.3	458	9 US-10-001-073-46	Sequence 46, Appli
17	8	2.3	461	9 US-10-225-567A-44	Sequence 44, Appli
18	8	2.3	462	9 US-10-001-073-44	Sequence 44, Appli
19	8	2.3	481	10 US-09-158-180-2	Sequence 2, Appli
20	8	2.3	501	9 US-10-244-367-44	Sequence 44, Appli
21	8	2.3	556	9 US-10-229-662-2	Sequence 2, Appli
22	8	2.3	556	10 US-09-795-691-2	Sequence 2, Appli
23	8	2.3	655	10 US-09-205-658-57	Sequence 57, Appli

24	8	2.3	655	10 US-09-844-353A-57	Sequence 57, Appli
25	8	2.3	728	9 US-09-836-392-21	Sequence 21, Appli
26	8	2.3	760	8 US-08-754-311B-2	Sequence 2, Appli
27	8	2.3	1237	9 US-10-108-605-211	Sequence 211, App
28	8	2.3	1394	9 US-10-108-605-213	Sequence 213, App
29	8	2.3	1987	9 US-10-132-382-6	Sequence 6, Appli
30	8	2.3	2013	9 US-10-132-382-2	Sequence 2, Appli
31	8	2.3	2014	9 US-10-132-382-8	Sequence 8, Appli
32	8	2.3	2040	9 US-10-132-382-4	Sequence 4, Appli
33	8	2.0	22	9 US-10-024-480-7	Sequence 7, Appli
34	7	2.0	23	9 US-10-311-111-29	Sequence 29, Appli
35	7	2.0	38	9 US-09-776-724A-165	Sequence 165, App
36	7	2.0	41	9 US-10-106-487-1	Sequence 1, Appli
37	7	2.0	45	10 US-09-864-761-41747	Sequence 41747, A
38	7	2.0	52	10 US-09-864-761-40439	Sequence 40439, A
39	7	2.0	75	9 US-10-106-658-6146	Sequence 6146, App
40	7	2.0	78	10 US-09-795-501-10	Sequence 10, Appli
41	7	2.0	88	9 US-09-738-626-6476	Sequence 6476, App
42	7	2.0	91	9 US-09-948-783-128	Sequence 128, App
43	7	2.0	92	9 US-09-892-877-127	Sequence 127, App
44	7	2.0	109	9 US-10-219-220-144	Sequence 144, App
45	7	2.0	109	10 US-09-864-761-47806	Sequence 47806, A

ALIGNMENTS

```
RESULT 1
US-10-188-947-4
; Sequence 4, Application US/10188947
; Publication No. US20030023993A1
; GENERAL INFORMATION:
; APPLICANT: MEDHITOV, Ruslan
; APPLICANT: HORNG, Tiffany
; APPLICANT: BARTON, Gregory
; TITLE OF INVENTION: TOLL/INTERLEUKIN-1 RECEPTOR ADAPTER PROTEIN (TIRAP)
; FILE REFERENCE: 044574-5101US
; CURRENT APPLICATION NUMBER: US/10/188,947
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/289,738
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: 60/289,815
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: 60/289,866
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: MISC_FEATURE
; OTHER INFORMATION: TIRAP polypeptide
US-10-188-947-4

Query Match          2.9%  Score 10; DB 9; Length 241;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      279 ASSSSVPASS 288
Db      2 ASSSSVPASS 11

RESULT 2
US-10-278-173-156
; Sequence 156, Application US/10278173
; Publication No. US20030061637A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Brown, Pierre
; APPLICANT: Riechmann, Jose-Luis
```

APPLICANT: Pineda, Omaira
APPLICANT: Zhang, James
APPLICANT: Yu, Guo-Liang
APPLICANT: Pilgrim, Marsha
APPLICANT: Keddle, James
APPLICANT: Heard, Jacqueline
APPLICANT: Reuber, Lynne
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc
APPLICANT: Samaha, Raymond
TITLE OF INVENTION: POLYNUCLEOTIDES FOR ROOT TRAIT ALTERATION
FILE REFERENCE: MBI-009
CURRENT APPLICATION NUMBER: US/10/278,173
CURRENT FILING DATE: 2002-10-21
PRIOR APPLICATION NUMBER: US/09/533,392
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/125,814
PRIOR FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 177
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 156
LENGTH: 369
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G215
US-10-278-173-156

Query Match 2.9%; Score 10; DB 9; Length 369;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SALAVAAAAA 19
|||||
DB 52 SALAVAAAAA 61

RESULT 3
US-09-864-761-38908
Sequence 38908, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO: 38908
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004633.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
US-09-864-761-38908

Query Match 2.3%; Score 8; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 AVAAAAA 20
|||||
DB 3 AVAAAAA 10

RESULT 4
US-09-864-761-43779
Sequence 43779, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30

```

US-10-156-761-13308
: Sequence 13308, Application US/10156761
: Publication No. US20030119018A1
:
: GENERAL INFORMATION:
:
: APPLICANT: OMURA, SATOSHI
: APPLICANT: IKEDA, HARUO
: APPLICANT: ISHIKAWA, JUN
: APPLICANT: HORIKAWA, HIROSHI
: APPLICANT: SHIBA, TADAYOSHI
: APPLICANT: SAKAKI, YOSHIYUKI
: APPLICANT: HATTORI, MASAHIRA
:
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDE
:
: FILE REFERENCE: 249-262
:
: CURRENT APPLICATION NUMBER: US/10/156,761
:
: CURRENT FILING DATE: 2002-05-29
:
: PRIOR APPLICATION NUMBER: JP 2001-204089
:
: PRIOR FILING DATE: 2001-05-30
:
: PRIOR APPLICATION NUMBER: JP 2001-272697
:
: PRIOR FILING DATE: 2001-08-02
:
: NUMBER OF SEQ ID NOS: 15109
:
: SEQ ID NO 13308
:
: LENGTH: 320
:
: TYPE: PRT
:

```

ORGANISM: Streptomyces avermitilis
US-10-156-761-13308

Query Match 2.3%: Score 8; DB 9; Length 320;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 112 AVSPADO 119
|||||||
Db 304 AVSPADO 311

RESULT 8
US-10-086-156-6
Sequence 6, Application US/10086156
Publication No. US20030054989A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING TWO NOVEL HUMAN POTASSIUM CHANNEL BETA-SUB
FILE REFERENCE: D0115NP
CURRENT APPLICATION NUMBER: US/10/086,156
CURRENT FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/272,190
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/274,258
PRIOR FILING DATE: 2001-03-07
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 338
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-10-086-156-6

Query Match 2.3%: Score 8; DB 9; Length 338;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 AVAAAAA 20
|||||||
Db 65 AVAAAAA 72

RESULT 9
US-10-071-458-4
Sequence 4, Application US/10071458
Publication No. US20030114371A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN POTASSIUM CHANNEL BETA-SUB
FILE REFERENCE: D0114.NP
CURRENT APPLICATION NUMBER: US/10/071,458
CURRENT FILING DATE: 2002-02-07
PRIOR APPLICATION NUMBER: US 60/267,039
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 60/281,224
PRIOR FILING DATE: 2001-04-03
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 338
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-10-071-458-4

Query Match 2.3%: Score 8; DB 9; Length 338;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 13 AVAAAAA 20
|||||||

Db 65 AVAAAAA 72

RESULT 10
US-10-286-264-112
Sequence 112, Application US/10286264
Publication No. US20030093837A1
GENERAL INFORMATION:
APPLICANT: Keddie, James
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Ratcliffe, Oliver
APPLICANT: Zhang, James
APPLICANT: Jjiang, Cai-Zhong
APPLICANT: Pineda, Omaira
APPLICANT: Heard, Jacqueline
APPLICANT: Yu, Guo-Liang
APPLICANT: Adam, Luc
APPLICANT: Broun, Pierie
APPLICANT: Reuber, Lynne
APPLICANT: Pilgrim, Marsha
APPLICANT: Samaha, Raymond
TITLE OF INVENTION: POLYNUCLEOTIDES FOR SEED TRAIT ALTERATION
FILE REFERENCE: MBI-008
CURRENT APPLICATION NUMBER: US/10/286,264
CURRENT FILING DATE: 2002-11-01
PRIOR APPLICATION NUMBER: 60/125,814
PRIOR FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 165
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 112
LENGTH: 374
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G670
US-10-286-264-112

Query Match 2.3%: Score 8; DB 9; Length 374;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 302 PSSSSATN 309
|||||||
Db 164 PSSSSATN 171

RESULT 11
US-09-934-455-304
Sequence 304, Application US/09934455
Publication No. US20030121070A1
GENERAL INFORMATION:
APPLICANT: Adam, Luc
APPLICANT: Creelman, Robert
APPLICANT: Dubell, Arnold
APPLICANT: Heard, Jacqueline
APPLICANT: Jjiang, Cai-Zhong
APPLICANT: Keddie, James
APPLICANT: Pilgrim, Marsha
APPLICANT: Ratcliffe, Oliver
APPLICANT: Reuber, Lynne
APPLICANT: Riechmann, Jose Luis
APPLICANT: Yu, Guo-Liang
APPLICANT: Pineda, Omaira
TITLE OF INVENTION: Genes for Modifying Plant Traits IV
FILE REFERENCE: MBI-0025
CURRENT APPLICATION NUMBER: US/09/934,455
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/227439
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: MBI-0022
PRIOR FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: MBI-0023
PRIOR FILING DATE: 2001-04-17

NUMBER OF SEQ ID NOS: 516
 SOFTWARE: Patentin version 3.1
 SEQ ID NO 304
 LENGTH: 374
 TYPE: PRT
 ORGANISM: Arabidopsis thaliana
 US-09-934-455-304

Query Match
 Best Local Similarity 100.0%; Score 8; DB 9; Length 374;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 PSSSSATN 309
 Db 164 PSSSSATN 171

RESULT 12
 US-10-156-761-12190
 Sequence 12190, Application US/10156761
 Publication No. US20030119018A1
 GENERAL INFORMATION:
 APPLICANT: OMURA, SATOSHI
 APPLICANT: IKEDA, HARUO
 APPLICANT: ISHIKAWA, JUN
 APPLICANT: HORIKAWA, HIROSHI
 APPLICANT: SHIBA, TADAYOSHI
 APPLICANT: SAKAKI, YOSHIYUKI
 APPLICANT: HATTORI, MASAHIRA
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-262
 CURRENT APPLICATION NUMBER: US/10/156,761
 PRIOR FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO 12190
 LENGTH: 427
 TYPE: PRT
 ORGANISM: Streptomyces avermitilis
 US-10-156-761-12190

Query Match
 Best Local Similarity 100.0%; Score 8; DB 9; Length 427;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 VAAAAAAL 21
 Db 241 VAAAAAAL 248

RESULT 13
 US-10-108-605-199
 Sequence 199, Application US/10108605
 Patent No. US20020160934A1
 GENERAL INFORMATION:
 APPLICANT: Broadus, Julie
 APPLICANT: Stam, Lynn
 APPLICANT: Bachmann, Jane
 APPLICANT: Kamdar, Kim
 TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
 TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
 FILE REFERENCE: 31133B
 CURRENT APPLICATION NUMBER: US/10/108,605
 PRIOR FILING DATE: 2002-03-27
 PRIOR APPLICATION NUMBER: US 09/761,142
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: US 60/176,418
 PRIOR FILING DATE: 2000-01-14
 NUMBER OF SEQ ID NOS: 361
 SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 199
 LENGTH: 440
 TYPE: PRT
 ORGANISM: Drosophila melanogaster
 US-10-108-605-199

Query Match
 Best Local Similarity 100.0%; Score 8; DB 9; Length 440;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 AVAAAAAA 20
 Db 325 AVAAAAAA 332

RESULT 14
 US-10-108-605-219
 Sequence 219, Application US/10108605
 Patent No. US20020160934A1
 GENERAL INFORMATION:
 APPLICANT: Broadus, Julie
 APPLICANT: Stam, Lynn
 APPLICANT: Bachmann, Jane
 APPLICANT: Kamdar, Kim
 TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
 TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
 FILE REFERENCE: 31133B
 CURRENT APPLICATION NUMBER: US/10/108,605
 PRIOR FILING DATE: 2002-03-27
 PRIOR APPLICATION NUMBER: US 09/761,142
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: US 60/176,418
 PRIOR FILING DATE: 2000-01-14
 NUMBER OF SEQ ID NOS: 361
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 219
 LENGTH: 440
 TYPE: PRT
 ORGANISM: Drosophila melanogaster
 US-10-108-605-219

Query Match
 Best Local Similarity 100.0%; Score 8; DB 9; Length 440;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 AVAAAAAA 20
 Db 325 AVAAAAAA 332

RESULT 15
 US-10-108-605-243
 Sequence 243, Application US/10108605
 Patent No. US20020160934A1
 GENERAL INFORMATION:
 APPLICANT: Broadus, Julie
 APPLICANT: Stam, Lynn
 APPLICANT: Bachmann, Jane
 APPLICANT: Kamdar, Kim
 TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
 TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
 FILE REFERENCE: 31133B
 CURRENT APPLICATION NUMBER: US/10/108,605
 PRIOR FILING DATE: 2002-03-27
 PRIOR APPLICATION NUMBER: US 09/761,142
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: US 60/176,418
 PRIOR FILING DATE: 2000-01-14
 NUMBER OF SEQ ID NOS: 361
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 243
 LENGTH: 440
 TYPE: PRT

; ORGANISM: Drosophila melanogaster
US-10-108-605-243

Query Match 2.3%; Score 8; DB 9; Length 440;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 AVAAAAA 20
| | | | | | |
Db 325 AVAAAAA 332

Search completed: July 11, 2003, 11:40:17
Job time : 1524 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 10:49:09 ; Search time 39 Seconds
(without alignments)
1192.422 Million cell updates/sec

Title: US-09-654-652A-3
Perfect score: 1824
Sequence: 1 MNIKKTAVKSAALVAATAA.....AKGAKVNPNGHKRYVNFEEH 349

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A.Geneseq_101002:.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	257.5	14.1	308	10	AAW93001
2	250.5	13.7	242	10	AAW93000
3	248	13.6	214	22	AAE07317
4	248	12.6	239	11	AAW06621
5	229	12.6	208	23	ABB76858
6	229	12.6	214	23	ABB76859
7	228	12.5	234	11	AAW03775
8	228	12.5	237	11	AAW05803
9	222.5	12.2	237	11	AAW06622
10	190.5	10.4	245	19	AAW37884

11	185	10.1	282	21	AAW48467	Arabidopsis thalia
12	185	10.1	282	23	ABB93737	Herbicidally activ
13	173	9.5	282	21	AAW32464	Arabidopsis thalia
14	173	9.5	282	23	ABB91456	Herbicidally activ
15	172.5	9.5	277	21	AAW52114	Arabidopsis thalia
16	172.5	9.5	277	23	ABB93169	Herbicidally activ
17	170.5	9.3	269	21	AAW18644	Arabidopsis thalia
18	170.5	9.3	282	21	AAW18643	Arabidopsis thalia
19	170.5	9.3	282	23	ABB93168	Herbicidally activ
20	170.5	9.3	282	21	AAW18642	Arabidopsis thalia
21	164	9.0	286	23	ABB93103	Herbicidally activ
22	161	8.8	284	23	ABB93853	Herbicidally activ
23	159.5	8.7	287	23	ABB93104	Herbicidally activ
24	157.5	8.6	269	21	AAW20347	Arabidopsis thalia
25	157.5	8.6	269	21	AAW20350	Arabidopsis thalia
26	157.5	8.6	269	21	AAW2653	Arabidopsis thalia
27	157.5	8.6	269	23	ABB93167	Herbicidally activ
28	157.5	8.6	290	21	AAW20346	Arabidopsis thalia
29	157.5	8.6	297	21	AAW2652	Arabidopsis thalia
30	157.5	8.6	300	21	AAW20349	Arabidopsis thalia
31	156.5	8.6	269	21	AAW20329	Arabidopsis thalia
32	156.5	8.6	299	21	AAW20328	Arabidopsis thalia
33	155	8.5	845	22	AAW98062	Clostridium perfr
34	154.5	8.5	284	23	ABB93854	Herbicidally activ
35	154.5	8.5	306	17	AAW97362	Oerskovia beta-1,3
36	154	8.4	247	18	AAW11593	Protein encoded by
37	154	8.4	247	20	AAW08308	A. thaliana merist
38	145.5	8.0	285	23	ABB93852	Herbicidally activ
39	145.5	8.0	307	21	AAW40361	Arabidopsis thalia
40	145.5	8.0	310	21	AAW40360	Arabidopsis thalia
41	145.5	8.0	310	23	ABB90942	Herbicidally activ
42	145.5	8.0	312	21	AAW40359	Arabidopsis thalia
43	145	7.9	263	18	AAW29455	Oerskovia xanthine
44	145	7.9	305	23	ABB91796	Herbicidally activ
45	144.5	7.9	307	21	AAW16385	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAW93001 standard; Protein: 308 AA.
ID AAW93001
AC AAW93001;
XX
XX
XX 19-MAY-1999 (first entry)
DT
XX
XX
DE B. alkalophilus beta-glucanase protein.
XX
XX
KW Beta-glucanase; enzyme; glucan; removal; lichenan; membrane;
KW equipment; food industry; brewing.
XX
XX
OS Bacillus alkalophilus.
XX
XX
PN DEL19732751-A1.
XX
XX
PD 04-FEB-1999.
XX
XX
PF 30-JUL-1997; 97DE-1032751.
XX
XX
PR 30-JUL-1997; 97DE-1032751.
XX
XX
PA (HENK) HENKEL KGAA.
XX
XX
PI Hillen W, Maurer K;
XX
XX
DR WPL: 1999-122161/11.
DR N-PSDB: AAW02912.
XX
XX
XX New Bacillus alkalophilus beta-glucanase protein and gene - useful
PT for removing glucan and/or lichenan from membranes in the brewing
PT industry

XX Claim 6; Page 5; 8pp; German.

PS This sequence represents a novel beta-glucanase isolated from *Bacillus*
CC *alcalophilus* DSM 9956. This enzyme is useful for removing glucan and/or
CC lichenan from membranes and equipment in the food industry, especially
CC the brewing industry.

XX Sequence 308 AA;

Query Match 14.1%; Score 257.5; DB 20; Length 308;
Best Local Similarity 36.5%; Pred. No. 2.1e-13;

Matches 62; Conservative 22; Mismatches 73; Indels 13; Gaps 6;

QY 29 DFGCAELTYLLEEVYGFKEARMKMAASGVSSMFLYONGSEIADGRPWVEVDIEVLGN 88

DB 96 EYKAGELRTNOFYGYGLFEVNMKPAKSTGVSSLETTY-TGFWMDNDPMDIDIEFLGKD 154

QY 89 PGSFOSNITIGKAGAOXTSEKHNHVASPADQAFHTYGLWTPNYRWTVDSGOEVRKTEGG 148

DB 155 TTRVQFNYFTNGVG---NNEHYHELGFDASESFTVAFEMPESISWYVNGELYYTA--- 208

QY 149 QVSHULTGT-OGLRNLMSSEA-AWVGOFDESKLPLFQFINWVYKVTTP 196

DB 209 -TENIPQTPKIMNMLPGIGVDGTVFGDEDPVYTEXDWV---RYTP 254

RESULT 2

AAP95000
ID AAP95000 standard; Protein: 242 AA.

XX AAP95000;

DT 12-FEB-1997 (first entry)

XX *Bacillus subtilis* lichenase.

XX Beer production; fermentation; barley; beta-glucan; hydrolysis;
KM lichenase.

OS *Bacillus subtilis* Y-25.

PN JP01067181-A.

PD 13-MAR-1989.

PF 08-SEP-1987; 87JP-0224615.

PR 08-SEP-1987; 87JP-0224615.

XX (ASAK) ASAH I BREWERIES KK.

XX WPI: 1989-119863/16.

DR N-PSDB; AAN95000.

XX Recombinant plasmid used in beer prodn. - obtd. by integrating
PT lichenase gene derived from *Bacillus subtilis*, into vector

XX Disclosure; Fig 4; 7pp; Japanese.

CC The lichenase gene from *Bacillus subtilis* Y-25 is used for
CC transforming *Bacillus* hosts so that they show increased lichenase
CC expression. The recombinant lichenase enzyme produced by the
CC transformants is useful in beer production for decomposing beta-
CC glucan from barley.

XX Sequence 242 AA;

Query Match 13.7%; Score 250.5; DB 10; Length 242;
Best Local Similarity 35.5%; Pred. No. 5.8e-13;

Matches 61; Conservative 25; Mismatches 75; Indels 11; Gaps 6;

QY 23 TTVSADFGCAELTYLLEEVYGFKEARMKMAASGVSSMFLYONGSEIADGRPWVEVDI 82

DB 80 TSPSYNKFDCGERSVQTYGYGLFEVNMKPAKNTGVSSFTYTGT---DGTPMDEIDI 136

QY 83 EVLGKNGFSQSNITIGKAGAOXTSEKHNHVASPADQAFHTYGLWTPNYRWTVDSGOEY 142

DB 137 EFLGKDTTKVQFNYFTNGAG---NHEKIVDLGFDANANAYHTVAFDQPNISKYVGO-L 192

QY 143 RKTGGGVSNLTGTGGLRFLMSSEA-AWVGOFDESKLPLFQFINWVYKVTTP 193

DB 193 KHTATNDIPTTPGK--IMNMLMGTCVDEWLGSITNGVN-PLVHYDWVRYTK 241

RESULT 3

AAP95000
ID AAP95000 standard; Protein: 214 AA.

XX AAP95000;

DT 06-NOV-2001 (first entry)

DE Barley recombinant, thermostable (1,3-1,4)-beta-glucanase enzyme.

XX Barley; transgenic barley malt; carbohydrate degrading enzyme; chicken;
KM (1,3-1,4)-beta-glucanase; nutritional value; foodstuff; thermostable.

OS Hordeum vulgare.

PN W0200159141-A2.

PD 16-AUG-2001.

PF 09-FEB-2001; 2001WO-US04222.

PR 10-FEB-2000; 2000US-0181473.

PR 09-NOV-2000; 2000US-0247126.

XX (UNIV) UNIV WASHINGTON STATE RES FOUND.

PI Von Wettstein D, Huang J, Horvath H;

DR WPI: 2001-497082/54.

XX New barley based foodstuff for animals, i.e. chicken comprising
PT recombinant carbohydrate degrading enzyme (1,3-1,4)-beta-glucanase -

PS Claim 8; Page 37-38; 43pp; English.

CC The present invention relates to a food stuff comprising barley feed and
CC transgenic barley malt where the transgenic barley malt comprises a
CC recombinant carbohydrate degrading enzyme comprising a (1,3-1,4)-beta-
CC glucanase. The food stuff is useful to feed animal, preferably chickens

CC Barley is cheaper than corn, the principal foodstuff of chickens.
CC Chickens cannot efficiently utilize barley as an energy source as they
CC do not possess a gut enzyme that depolymerises beta-D-glucan, major
CC carbohydrate present in the barley endosperm. The invention provides a
CC barley based foodstuff which comprises a recombinant carbohydrate
CC degrading enzyme improving the nutritional value of the foodstuff.

CC The present sequence is barley recombinant, thermostable
CC (1,3-1,4)-beta-glucanase enzyme.

XX Sequence 214 AA;

Query Match 13.6%; Score 248; DB 22; Length 214;
Best Local Similarity 35.4%; Pred. No. 8e-13;

Matches 62; Conservative 27; Mismatches 70; Indels 16; Gaps 8;

QY 23 TTVSADFGCAELTYLLEEVYGFKEARMKMAASGVSSMFLYONGSEIADGRPWVEVDI 82

DB 52 TSPSYNKFDCGERSVQTYGYGLFEVNMKPAKNTGVSSFTYTGTPE---GTPMDEIDI 108

QY 83 EVLGKNGFSQSNITIGKAGAOXTSEKHNHVASPADQAFHTYGLWTPNYRWTVDSGOEY 142

DB 109 EFLGKDTTKVQFNYFTNGAGH---EKVISTLGFDAKSGPHTVAFDQPNISKYVGO-VL 164

OY	143	RKTGGGVSNLTGTGOG-LRFNLMWSESA-AVWGQDESKLPFGFIRINWKKYYKT	195
Dd	165	KHT---ATAPIPSIPGKIMMLMNMGTGYDDMLGSTNGAN-PLYAETDWMY---KYT	212
 RESULT 4 AA006621			
ID	AA006621	standard; protein; 239 AA.	
XX	AC	AA006621;	
XX	DT	09-JAN-1991 (first entry)	
XX	DE	Hybrid (1,3-1,4)-pre-beta-glucanase.	
XX	KW	Hybrid pre-beta-glucanase; glucans; beer; animal feed; poultry.	
XX	OS	Bacillus amyloliquefaciens, Bacillus macerans.	
XX	FT	Key	
XX	FT	Domain	
XX	FT	Domain	
XX	FT	Domain	
PX	PN	WO9009436-A.	
PX	PD	23-AUG-1990.	
PE	PE	16-FEB-1990; 90WO-DK00044.	
PX	PR	04-AUG-1989; 89DK-0003848.	
PR	PR	16-FEB-1989; 89DD-0325800.	
XX	PA	(CARL-) CARLSBERG A/S.	
XX	PA	(DEAK) AKAD WISSENSCHAFT DDR.	
XX	PI	Borriess R, Hofemeister J, Thomsen KK, Olsen O, Vonwettstein D;	
XX	DR	WPI: 1990-275129/36.	
XX	N-	N-PSDB; AA005832.	
PT	PT	New thermostable (1,3-1,4)-beta-glucanase - prep'd. using hybrid	
PT	PT	gene obt'd. using Bacillus amyloliquefaciens and B.macerans genes	
XX	PS	Disclosure: page 26; 84pp; English.	
XX	CC	This hybrid protein is encoded by the beta-glucanase-HI gene.	
XX	CC	Following processing of the signal peptide the mature protein	
XX	CC	is produced, comprising the amino terminus of the amylolique-	
XX	CC	factiens beta-glucanase and the carboxyl-terminal half of the	
XX	CC	B.macerans beta-glucanase. This hybrid protein is thermostable	
XX	CC	and hydrolyses beta-glycosidic linkages in (1,3-1,4)-beta-glucans.	
XX	CC	Reducing sugars are obtd. at high temps. and thus this enzyme can	
XX	CC	be used in the mfr. of food prodcs., esp. beer and animal feed (eg	
XX	CC	for feeding poultry). See also AA005833.	
SQ	Sequence	239 AA;	
OY	Query Match	13.6%; Score 248; DB 11; Length 239;	
OY	Best Local Similarity	35.4%; Pred. No. 9.3e-13;	
OY	Matches	62; Conservative 27; Mismatches 70; Indels 16; Gaps 8;	
Dd	23	TNVASAKDSSGAELITLEVQCGKRPARYKMAAAGSTVSMPLOYGSELADGRPVEVDI	82
Dd	77	TSPIYNKDCGENSVQYTGYLEVRMKRAKNNTGIASSFFTYGTE--GTPWDEIDI	133
OY	83	EVLGNKPSFOSNIITTKAGAKTSEKHNAVSPADAQFHHTYGLEMTPNYVFWTVDOGEV	142
Dd	134	EFLGIDTTFKVGFNYTTNGVGSH---EKVISLGFDASKRPHHYAFPMOIGYIKMYDG-VL	189
OY	143	RKTGGGVSNLTGTGOG-LRFNLMWSESA-AVWGQDESKLPFGFIRINWKKYYKT	195

```

Db      190 KHT--ATANIPSPFGKIMMLNMNGTGVDDWDLGSGYNGAN-PLVAYEYD---KIT 237

RESULT 5
ID      ABB76858
       ABB76858 standard; Protein: 208 AA.
XX      ABB76858;
AC      ABB76858;
DT      27-JUN-2002 (first entry)
DE      Bacterial Lichenase #1.
KW      Soil bacterium; enzyme: lichenase; endo-1,3/1,4-beta-glucanase; brewing.
XX      unidentified.
OS      WO200212511-A1.
XX      PN
XX      PD      14-FEB-2002.
XX      PF      27-JUL-2001; 2001WC-ES00303.
XX      PR      28-JUL-2000; 2000ES-0001922.
PA      (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
PI      Rodriguez Munoz V, Perez Mellado R;
DR      WPI: 2002-217195/27.
XX      N-PDB: AB15374.
PT      Producing nucleic acid encoding lichenase, useful for producing enzymes
PR      for improving filtration in brewing, comprises isolating the nucleic
PI      acid from soil bacteria -
PS      Claim 18; Page 18-19; 27pp; Spanish.
CC      The present sequence is a protein sequence from a soil bacterium for an
CC      enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)
CC      activity. The enzyme is useful in brewing and for degrading
CC      beta-glucanases that cause problems during filtration.
CX      CC
SQ      Sequence 208 AA;

Query Match          12.6%; Score 229; DB 23; Length 208;
Best Local Similarity 36.1%; Pred. No. 3e-11; Mismatches 67; Indels 10; Gaps 5;
Matches 56; Conservative 22;

QY      23 TNVSAKDSCAGELVTLEEVQYGFAPARKMAASGTVSMPFYONGSEIADGRPWVEVDI 82
       I : I : : : : I : I : I : I : I : I : I : I : I : I : I : I :
Db      55 TSPSVNRKDCCEENRVRQTYGYGLYVRKKRPKNKTICIVSFTTYGTFTE---GTPEDEIDI 111
       I : I : : : : I : I : I : I : I : I : I : I : I : I : I : I :

QY      83 EVLGNKPGSFCSNIITGKAGAOKTSEKHAAVSPAADOAFHTYGLEWTPNRYRWTVDGQEV 142
       I : I : : : : I : I : I : I : I : I : I : I : I : I : I : I :
Db      112 EFLGKDTKYVFNYTYTGAG---NHKEKLADI GFDAANAHHYAEPDMCPNSIKWYDQG-L 167
       I : I : : : : I : I : I : I : I : I : I : I : I : I : I : I :

QY      143 RTBEGQVSNLTGTQGLRFNLMSSEA-AWYGQFD 176
       I : I : : : : I : I : I : I : I : I : I : I : I : I : I : I :
Db      168 KHTATTQTIPAPRK--IMMNLINMGTDGVDWDLGSTYN 200
       I : I : : : : I : I : I : I : I : I : I : I : I : I : I : I :

RESULT 6
ABB76859
ID      ABB76859 standard; Protein: 214 AA.
XX      ABB76859;
AC      ABB76859;
DT      27-JUN-2002 (first entry)
DE      Bacterial Lichenase #2.
XX
```

XX	Soil bacterium; enzyme; lichenase; endo-1,3/1,4-beta-glucanase; brewing.
XX	unidentified.
OS	
PN	W0200212511-A1.
XX	
PD	14-FEB-2002.
XX	
PF	27-JUL-2001; 2001WO-ES00303.
XX	
PR	28-JUL-2000; 2000ES-0001922.
XX	
PA	(CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX	
PI	Rodriguez Munoz V, Perez Mellado R;
XX	
DR	WPI: 2002-217195/27.
N-PSDB; ABL53375.	
XX	
P	producing nucleic acid encoding lichenase, useful for producing enzymes
PT	for improving filtration in brewing, comprises isolating the nucleic
PT	acid from soil bacteria .
XX	
PS	Claim 18; Page 21-22; 27pp; Spanish.
CC	The present sequence is a protein sequence from a soil bacterium for an
CC	enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)
CC	activity. The enzyme is useful in brewing and for degrading
CC	beta-glucanases that cause problems during filtration.
SO	
Sequence	214 AA;
Query Match	12.6%; Score 229; DB 23; Length 214;
Best Local Similarity	36.1%; Pred. No. 3.2e-11;
Matches	56; Conservative 22; Mismatches 67; Indels 10; Gaps 5;
DQ	23 TNVSAKDFSGAELTYLIEVOYGFKEARMKMAASGTSVMFLYONGSEIADGRPVVEYDI 82 : : : : : : 60 TSPSYNKFDCGENSNVQTYGTGLYEVRKKPKAKNTGYSSFTTGTPE--GTPWDEIDI 116
Dd	83 EVLGKNPGSFGSNITTKRAGAOKTSEKHNAVSPPADQAFTHTGLEMTFENYYRWTVDGOEV 142 : : : : 117 EFLKRDFTKKVGFNYVTTCAG---NHKLADI GFDAANAYHTYAFDMQFNISIKWYDQG-L 172
OY	143 KRTGCGOVSNLTGTQGLRFNMSSESA-AMWGQEP 176 : : : : : 173 KHATFTQTIPAAPGK--IMMNLNGTGVDWLGSYN 205
Dd	
RESULT 7	
AAR03775	
ID	AAR03775 standard; protein: 234 AA.
XX	
AA	AAR03775;
XX	
DT	31-JUL-1990 (first entry)
XX	
DE	Thermostable beta-glucanase.
XX	
KW	Thermostable beta-glucanase; ss; Bacillus subtilis.
XX	
OS	Bacillus macerans.
XX	
PN	DD272102-A.
XX	
PD	27-SEP-1989.
XX	
PF	12-MAY-1988; 88DD-0315706.
XX	
PR	12-MAY-1988; 88DD-0315706.
XX	
PA	(DEAK) AKAD WISSENSCHAFT DDR.
XX	

```

PI   Borriss R. 1990-067913/10.
DR   WPI: 1990-067913/10.
DR   N-PSDB: AAO03519.
XX
XX   The most stable beta-glucanase production -
PT   using Bacillus subtilis transformed with gene from Bacillus
PT   macerans.
XX
XX   Fig 1 : 9pp; German.
XX
XX   The gene encoding Bacillus macerans thermostable
CC   beta-glucanase is expressed in Bacillus subtilis. The
CC   enzyme is useful for lowering the viscosity of brewing
CC   mashes and in the production of feedstuff.
XX
XX   Sequence 234 AA:
SQ
XX
XX   Query Match 12.5%, Score 228; DB 11; Length 234;
XX   Best Local Similarity 34.3%; Pred. No. 4.3e-11;
XX   Matches 60; Conservative 25; Mismatches 74; Indels 16; Gaps 8;
QY
QY   23 TNSAKPFSSGALTYLLEVOYGRKREARKKMAAGTSSMFLYONGSLADGRPMVEVDI 82
Db   72 TSSAYNKFDCAEYRSTNIYIGLYEVSKKPAKNTGIVSSFTYGP--AHGTQWDEIDI 128
QY   83 EYLGRNPGSEFOSNITITGKAGAKTSEKHHAVSPADQAFTHTYGLEWTPNYYRTVDSQEV 142
Db   129 EFLKRDPTTKVQENYYTNGVGH---EKVISLGFDSKGFHHTYARDMDQGYIKWYVDG-VL 184
QY   143 RKTGGGVSNLTGQG-LRFLNLSESA-AAVGQFDESKLPLFOFINWKKYKXTT 195
Db   185 KHT--ATANIPSTPGKIMMLMNGTGVDWLGSYNGAN-PLYAEYDWV--KYYT 232

RESULT 8
AAR05803
AAR05803 standard; protein: 237 AA.
XX
XX   AAR05803;
AC
XX
XX   08-NOV-1990 (first entry)
DT
XX
XX   Heat-stable endo-beta-1,3-1,4-glucanase.
DE
XX
XX   Endo-beta-1,3-1,4-glucanase; barley; brewing.
KW
XX
XX   Bacillus macerans.
OS
XX
XX   DD275704-A.
PN
XX
XX   31-JAN-1990.
PD
XX
XX   23-SEP-1988; 88DD-0320082.
PE
XX
XX   23-SEP-1988; 88DD-0320082.
PR
XX
XX   (DEAK ) AKAD WISSENSCHAFT DDR.
XX
XX   Borriss R, Wobus U, Mendel R-R, Baumlein H;
PI
XX
XX   WPI: 1990-210631/28.
DR
XX
XX   N-PSDB: AAO05167.
XX
XX   Piepu. of barley plants expressing heat stable beta-glucanase -
PT   by transforming cells with appropriate vector then regeneration
PT   giving seeds useful in brewing without conversion to malt.
XX
XX   Disclosure: ; p; German.
PS
XX
XX   The corresp. gene and the plant promoter region are inserted into an
CC   expression vector and used to transform barley cells. The transformants
CC   are used to regenerate barley plants which are useful in brewing. This
CC

```


CC of grain (e.g. barley or wheat) in the brewing and fermentation
CC Industries to increase carbon substrate availability and to maximise
CC production of desired products.

SO Sequence 245 AA;

Query Match 10.4%; Score 190.5; DB 19; Length 245;

Best Local Similarity 31.6%; Pred. No. 6.5e-08;

Matches 54; Conservative 17; Mismatches 87; Indels 13; Gaps 6;

OY 21 LTTNVSAGDFSGALTYLEVOYKREARKMAAASCTVSMPLVYONGSEIADGRPVEV 80

DB 79 LITDRGSGYTCGEYRTKNTYGYGMFOVNMKPIKPNQVVSFFTYTGPS--DGTKWDEI 135

OY 81 DIEVLGNPGSPSNITITGKAGAKTSEKHNAVSPADDAFHTYGLEMTPNYKWTYDQ 140

DB 136 DIEFLGDTTKVQPNYYTNGSGHH--EHITHLGFDPSOGFHYGFWFARNSTIWTYDGT 192

OY 141 EVKRTGCGVSNLTGTG-LRFNLMSESA-AWVGQFDESKLPLFOFINMV 189

DB 193 AVYTA---YDNIPDPFGKIMNANWNGIGVDWLRPFN-GRNISAVIDMV 238

RESULT 11

AAG48467

ID AAG48467 standard; Protein: 282 AA.

AC AAG48467;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 61207.

KM Protein identification: signal transduction pathway: metabolic pathway;

KM hybridisation assay: genetic mapping; gene expression control; promoter;

KM termination sequence.

OS Arabidopsis thaliana.

OS EP1033405-A2.

PN 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

PE 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0123788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129645.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132488.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135529.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139453.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139753.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139819.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0143977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR	08-JUL-1999	9905-01453951
PR	02-AUG-1999	9905-0146388
PR	02-AUG-1999	9905-0146388
PR	02-AUG-1999	9905-0146389
PR	04-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147260
PR	06-AUG-1999	9905-0147300
PR	06-AUG-1999	9905-0147416
PR	09-AUG-1999	9905-0147493
PR	09-AUG-1999	9905-0147935
PR	10-AUG-1999	9905-0147935
PR	10-AUG-1999	9905-0148171
PR	11-AUG-1999	9905-0148319
PR	12-AUG-1999	9905-0148341
PR	13-AUG-1999	9905-0148565
PR	13-AUG-1999	9905-0148684
PR	16-AUG-1999	9905-0149368
PR	17-AUG-1999	9905-0149175
PR	18-AUG-1999	9905-0149426
PR	18-AUG-1999	9905-0149426
PR	20-AUG-1999	9905-0149722
PR	20-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149825
PR	23-AUG-1999	9905-0149902
PR	23-AUG-1999	9905-0149930
PR	25-AUG-1999	9905-0150566
PR	26-AUG-1999	9905-0150584
PR	27-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151080
PR	30-AUG-1999	9905-0151303
PR	31-AUG-1999	9905-0151438
PR	01-SEP-1999	9905-0151930
PR	07-SEP-1999	9905-0153363
PR	10-SEP-1999	9905-0153707
PR	13-SEP-1999	9905-0153758
PR	13-SEP-1999	9905-0154018
PR	16-SEP-1999	9905-0154039
PR	20-SEP-1999	9905-0154779
PR	22-SEP-1999	9905-0155139
PR	23-SEP-1999	9905-0155486
PR	24-SEP-1999	9905-0155486
PR	28-SEP-1999	9905-0156458
PR	29-SEP-1999	9905-0156559
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157573
PR	06-OCT-1999	9905-0157865
PR	07-OCT-1999	9905-0158029
PR	08-OCT-1999	9905-0158232
PR	12-OCT-1999	9905-0158365
PR	13-OCT-1999	9905-0159293
PR	13-OCT-1999	9905-0159294
PR	14-OCT-1999	9905-0159329
PR	14-OCT-1999	9905-0159330
PR	14-OCT-1999	9905-0159330
PR	14-OCT-1999	9905-0159331
PR	14-OCT-1999	9905-0159637
PR	14-OCT-1999	9905-0159638
PR	18-OCT-1999	9905-0159584
PR	21-OCT-1999	9905-0160741
PR	21-OCT-1999	9905-0160767
PR	21-OCT-1999	9905-0160768
PR	21-OCT-1999	9905-0160770
PR	21-OCT-1999	9905-0160814
PR	22-OCT-1999	9905-0160815
PR	22-OCT-1999	9905-0160980
PR	22-OCT-1999	9905-0160981
PR	22-OCT-1999	9905-0160981
PR	25-OCT-1999	9905-0161404
PR	25-OCT-1999	9905-0161405
PR	25-OCT-1999	9905-0161406
PR	26-OCT-1999	9905-0161359

	Query Match	Similarity	10.1%	Score 185	DB 21	Length 282
PR	26-OCT-1999;	99US-0161360.				
PR	26-OCT-1999;	99US-0161361.				
PR	28-OCT-1999;	99US-0161920.				
PR	28-OCT-1999;	99US-016192.				
PR	28-OCT-1999;	99US-0161993.				
PR	29-OCT-1999;	99US-0162142.				
Query Match	Similarity	27.4%	Pred. No. 2.3e-07.			
Matches	59;	Conservative	34;	Mismatches	98;	Indels 24; Gaps 9
OY	24 NVSAKDFEGALYETLEVOYGKFEARMMAA--ASGTYSMFLYONGSEIADGRPWVEVD	81				
Db	53 SLSDLKFGSGQSHQEFLYGVYKMLVPEGNSAGVTYTFYLK-----SPGTWIDEID	106				
OY	82 IEVLGK--NGSFDOSNIITGKAGQKISEKHNAVSPAADQAFHTYGLETPNTNYRWIVD	138				
Db	107 FEEFGNISGHPEYTLHTNYVT-KGTGDKKQGFHMLPDPYVD--FHTYCIIMNPFQRFVFTYD	163				
OY	139 GOEYKRTGGGVSNL--TGTOGLRF--NLMSSESAWVG---QPEESKLPLFOFINWVKV	191				
Db	164 GIPREFNSNELSVPPFKHQPMRIYASLMEMEHATRGLEKTDMSKAPFTAFIRNNV	223				
OY	192 KYTPGCGEGSGDFTLDWTFNDFTFDGS--RWGK	223				
Db	224 DACVWSNGKSSCSANSSWFTQVLDKGRNRVKAQ	258				
RESULT 12						
ABB93737						
ID	ABB93737 standard; Protein: 282 AA.					
XX	ABB93737;					
XX						
DT	31-MAY-2002 (first entry)					
XX						
DE	Herbicideally active polypeptide SEQ ID NO 2948.					
XX						
KW	Herbicideal; plant; agriculture; herbicide.					
XX						
OS	Arabidopsis thaliana.					
XX						
PN	WO200210210-A2.					
XX						
PD	07-FEB-2002.					
XX						
PF	28-AUG-2001; 2001WO-EP09892.					
XX						
PR	28-AUG-2001; 2001WO-EP09892.					
XX						
PA	(FARB) BAYER AG.					
XX						
PI	Tietjen K, Weidler M;					
XX						
DR	WPI: 2002-269010/31.					
PT	Identifying plant target proteins for herbicideally active compounds,					
PT	comprising aligning and comparing nucleic acid or amino acid sequences					
PT	from plant with nucleic acid or amino acid sequences from non-plant					
PT	organisms -					
XX						
PS	Claim 5; SEQ ID NO 2948; 261pp + Sequence Listing; English.					
XX						
CC	The invention relates to identifying target proteins					
CC	(ABB93737-ABB94016) for herbicideally active compounds, comprising					
CC	aligning and comparing nucleic acid or amino acid sequences from plant					
CC	with nucleic acid or amino acid sequences from non-plant organisms using					
CC	suitable search parameters, where plant sequences having an E-value					
CC	greater by a factor of 3 than the E-value of most similar non-plant					
CC	sequences are selected. The polypeptides or nucleic acids encoding them					
CC	are useful for identifying modulators. The identified modulators are					
XX	useful as herbicides.					

Seq	Sequence	282 AA	10.1%	Score 185	DB 23	Length 282
Query Match			27.4%	Pred. NO. 2.3e-07		
Best Local Similarity			34	Mismatches 98	Indels 24	Gaps 9
Matches						
Db	24	NVSAKDFGAGELTYLLEEVQYGRFARMKMAA--ASGTSSMFLYNGSEIADGRWPEVD	81			
Db	53	SLSDKFEFGSGFQSHQEFRLYGVKVEQMKLVPGNSAGVATYTLK-----SGCTTWDEID	106			
Qy	82	IEVLGK---NGSFQSNITITGKAGAKQKSEKHNAAVSPAADOAFHTYGLEMTPNVYRMTVD	138			
Db	107	FEFLGNISGHYTLHTNYVT-KGTGDKRQOQHLMFDPYVD--FHYYCIIMNPQRVITFID	163			
Db	139	GOEVRKTEGGGVSNL--TGTGGLRF--NLMSSESAAWVG---QFDESKLPLEQFINWYV	191			
Qy	164	GIPRIREFKNSBALGVPPFKHQPMLRYASLMEAEHMTATGKLEKTDMSKAPFTAFYRNVN	223			
Qy	192	YKTYPGGEGGSDFTLMDTFDFDGS---RWGK	223			
Db	224	DACVMSNGKSSCSANSSWFTGVLDLFCKGRNVKMAQ	258			
RESULT 13						
AAG32464						
ID	AAG32464	standard; Protein: 282 AA.				
XX						
AC	AAG32464:					
XX						
DE	17-OCT-2000	(first entry)				
XX						
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 39167.					
KW	Protein identification; signal transduction pathway; metabolic pathway;					
KM	hybridisation assay; genetic mapping; gene expression control; promoter;					
XX	termination sequence.					
OS	Arabidopsis thaliana.					
XX						
PN	EP1033405-A2.					
XX						
PD	06-SEP-2000.					
XX						
PF	25-FEB-2000; 2000EP-0301439.					
XX						
XX	25-FEB-1999; 99US-0121825.					
PR	05-MAR-1999; 99US-0123180.					
PR	09-MAR-1999; 99US-0123548.					
PR	23-MAR-1999; 99US-0125788.					
PR	23-MAR-1999; 99US-0126264.					
PR	29-MAR-1999; 99US-0126785.					
PR	01-APR-1999; 99US-0127462.					
PR	06-APR-1999; 99US-0128234.					
PR	08-APR-1999; 99US-0128714.					
PR	16-APR-1999; 99US-0129845.					
PR	19-APR-1999; 99US-0130077.					
PR	21-APR-1999; 99US-0130449.					
PR	23-APR-1999; 99US-0130510.					
PR	23-APR-1999; 99US-0130891.					
PR	28-APR-1999; 99US-0131449.					
PR	30-APR-1999; 99US-0132048.					
PR	30-APR-1999; 99US-0132407.					
PR	04-MAY-1999; 99US-0132484.					
PR	05-MAY-1999; 99US-0132485.					
PR	06-MAY-1999; 99US-0132486.					
PR	06-MAY-1999; 99US-0132487.					
PR	07-MAY-1999; 99US-0132863.					
PR	11-MAY-1999; 99US-0134256.					
PR	14-MAY-1999; 99US-0134218.					
PR	14-MAY-1999; 99US-0134219.					
PR	14-MAY-1999; 99US-0134221.					
PR	14-MAY-1999; 99US-0134370.					
PR	16-MAY-1999; 99US-0134768.					

PR	19-MAY-1999	9905-0134544
PR	20-MAY-1999	9905-0135124
PR	21-MAY-1999	9905-0135353
PR	22-MAY-1999	9905-0135629
PR	23-MAY-1999	9905-0136061
PR	24-MAY-1999	9905-0136592
PR	27-MAY-1999	9905-0136782
PR	28-MAY-1999	9905-0136782
PR	01-JUN-1999	9905-0137222
PR	03-JUN-1999	9905-0137528
PR	04-JUN-1999	9905-0137502
PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138547
PR	10-JUN-1999	9905-0138847
PR	14-JUN-1999	9905-0139119
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139453
PR	17-JUN-1999	9905-0139492
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139459
PR	18-JUN-1999	9905-0139460
PR	18-JUN-1999	9905-0139461
PR	18-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0140033
PR	23-JUN-1999	9905-0140354
PR	23-JUN-1999	9905-0140355
PR	24-JUN-1999	9905-0140695
PR	28-JUN-1999	9905-0140823
PR	29-JUN-1999	9905-0140991
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141847
PR	01-JUL-1999	9905-0142154
PR	02-JUL-1999	9905-0142051
PR	06-JUL-1999	9905-0142390
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142820
PR	12-JUL-1999	9905-0143277
PR	13-JUL-1999	9905-0143342
PR	14-JUL-1999	9905-0143624
PR	15-JUL-1999	9905-0144005
PR	16-JUL-1999	9905-0144085
PR	16-JUL-1999	9905-0144086
PR	16-JUL-1999	9905-0144225
PR	19-JUL-1999	9905-0144331
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144333
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144335
PR	20-JUL-1999	9905-0144352
PR	20-JUL-1999	9905-0144632
PR	20-JUL-1999	9905-0144684
PR	21-JUL-1999	9905-0144814
PR	21-JUL-1999	9905-0144816
PR	21-JUL-1999	9905-0145088
PR	21-JUL-1999	9905-0145085
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145192
PR	23-JUL-1999	9905-0145145
PR	23-JUL-1999	9905-0145218
PR	23-JUL-1999	9905-0145224
PR	26-JUL-1999	9905-0145576
PR	27-JUL-1999	9905-0145513
PR	27-JUL-1999	9905-0145518
PR	27-JUL-1999	9905-0145519

```

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147418.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

```

```

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 9.5% Score 173; DB 21; Length 282;
 Best Local Similarity 25.6%; Pred. No. 2.3e-06;
 Matches 64; Conservative 36; Mismatches 92; Indels 58; Gaps 12;

```

QY 24 NVSANDPSGADLYLLEVOYGFKFAKMAA--ASGVSSMFLYONGSEIADGRWFEVD 81
  ::||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 53 SLSDKSSGSGSPQSNQEFYLGKAEVOMKLVPGNSAGVTYTYLK-----SPGTTWDEID 106
  ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 82 IEVLCK---HPSRQSNITITKAGAKQKTESEHHAVSPADDAFHYGLEMTPNVRYRYVD 138
  ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 107 FEELGNISGHYYTLTNYVT-KGTGDKDOQFLWPDPTVN--FHYTCITMNPQRIITFYVD 163
  |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 139 GOEVRKTEGCVSNL--TGTOGLRF--NLMSESAWVG---QFDESKLPLEQFINWVKV 191
  |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 164 GIPREFKNPRLAIGVPPFTROPMLRYASLMEAEHMTATRGLEKTMDSKAPFTAF----- 217
  |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 192 KYRTPGQEGKGSDFLTLDNEDTPDGSRMGKGMWTFDGNRYDLTKNIYSRDGMLIAL 251
  |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 218 YR-----NINV-DGCVMAWGKSSCSANSPWFTQK-----L 246
  |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 252 TRKGOSEFNG 261
  |||
DB 247 DSGQTRMKG 256

```

RESULT 14

ABB91456 ID ABB91456 standard; Protein; 282 AA.

XX ABB91456;

XX 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 667.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP09892.

XX 28-AUG-2001; 2001WO-EP09892.

XX (FARB) BAYER AG.

XX Tietjen K, Weidner M.

XX WPI; 2002-269010/31.

Identifying plant target proteins for herbicidally active compounds,
 comprising aligning and comparing nucleic acid or amino acid sequences
 from plant with nucleic acid or amino acid sequences from non-plant
 organisms -

Claim 5; SEQ ID NO 667; 261pp + Sequence Listing; English.

The invention relates to identifying target proteins
 (ABB90790-ABB94016) for herbicidally active compounds, comprising
 aligning and comparing nucleic acid or amino acid sequences from plant
 with nucleic acid or amino acid sequences from non-plant organisms using
 suitable search parameters, where plant sequences having an E-value
 greater by a factor of 3 than the E-value of most similar non-plant

CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
XX useful as herbicides.

SO Sequence 282 AA:

Query Match 9.5%; Score 173; DB 23; Length 282;

Best Local Similarity 25.6%; Pred. No. 2.3e-06;
Matches 64; Conservative 36; Mismatches 92; Indels 58; Gaps 12;

```
QY 24 NVSAKDFSGALYILEEYQYGFARMAA--ASGYVSFPLQNGSEIDGRFWEVD 81
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 53 SLSDKSSGSGFQSGFELGYKAEYQMKLVPGNSAGYVTTFTYK-----SPGTTWDEID 106
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 82 IEVLGK---NPGSFQSNITTGAKAQKTEKHNHAPDAQFHYGLETPVNYRWTVVD 138
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 107 FEFGLNIGHRYTLHTNYIT- KGTGDKEQOFLHNFDPVNV--FHTYCTTWMPQRIETVVD 163
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 139 GQEVKRTGEGGVSNL--TGTGLRF--NLWSESAWVG---QFDESKLPLEOFINWVKV 191
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 164 GIPREFKNPRAIGVPEPTROPMLRYASLWEAEHMAHTRGLEKTDWMSKAPPTAF----- 217
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 192 YKTPGQEGGSDTFLDWTQNFDFDGSRMGKGMTEFGNVDLTDKNITVRDGLIAL 251
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 218 YR-----NYNV-DGCVWANGSSCSANSPWFTQK-----L 246
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 252 TRKGQSEFNG 261
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 247 DSNQOTRMKG 256
```

RESULT 15

AGS2114
ID AGS2114 standard; Protein: 277 AA.

XX AGS2114;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 66212.

KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123160.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 16-APR-1999; 99US-0128714.

XX 19-APR-1999; 99US-0129845.

XX 21-APR-1999; 99US-0130077.

XX 23-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132464.

XX 06-MAY-1999; 99US-0132465.

XX 06-MAY-1999; 99US-0132486.

```
PR 06-MAY-1999; 99US-0132487.
```

```
PR 07-MAY-1999; 99US-0132863.
```

```
PR 11-MAY-1999; 99US-0134256.
```

```
PR 14-MAY-1999; 99US-0134218.
```

```
PR 14-MAY-1999; 99US-0134219.
```

```
PR 14-MAY-1999; 99US-0134221.
```

```
PR 14-MAY-1999; 99US-0134370.
```

```
PR 18-MAY-1999; 99US-0134376.
```

```
PR 19-MAY-1999; 99US-0134941.
```

```
PR 20-MAY-1999; 99US-0135124.
```

```
PR 21-MAY-1999; 99US-0135353.
```

```
PR 24-MAY-1999; 99US-0135629.
```

```
PR 25-MAY-1999; 99US-0136021.
```

```
PR 27-MAY-1999; 99US-0136392.
```

```
PR 28-MAY-1999; 99US-0136782.
```

```
PR 01-JUN-1999; 99US-0137222.
```

```
PR 03-JUN-1999; 99US-0137528.
```

```
PR 04-JUN-1999; 99US-0137502.
```

```
PR 07-JUN-1999; 99US-0137724.
```

```
PR 08-JUN-1999; 99US-0138094.
```

```
PR 10-JUN-1999; 99US-0138540.
```

```
PR 10-JUN-1999; 99US-0138847.
```

```
PR 14-JUN-1999; 99US-0139119.
```

```
PR 16-JUN-1999; 99US-0139452.
```

```
PR 16-JUN-1999; 99US-0139453.
```

```
PR 17-JUN-1999; 99US-0139492.
```

```
PR 18-JUN-1999; 99US-0139454.
```

```
PR 18-JUN-1999; 99US-0139455.
```

```
PR 18-JUN-1999; 99US-0139456.
```

```
PR 18-JUN-1999; 99US-0139457.
```

```
PR 18-JUN-1999; 99US-0139458.
```

```
PR 18-JUN-1999; 99US-0139459.
```

```
PR 18-JUN-1999; 99US-0139460.
```

```
PR 18-JUN-1999; 99US-0139461.
```

```
PR 18-JUN-1999; 99US-0139462.
```

```
PR 18-JUN-1999; 99US-0139463.
```

```
PR 18-JUN-1999; 99US-0139750.
```

```
PR 18-JUN-1999; 99US-0139763.
```

```
PR 21-JUN-1999; 99US-0139817.
```

```
PR 22-JUN-1999; 99US-0139859.
```

```
PR 23-JUN-1999; 99US-0140353.
```

```
PR 23-JUN-1999; 99US-0140354.
```

```
PR 24-JUN-1999; 99US-0140695.
```

```
PR 28-JUN-1999; 99US-0140823.
```

```
PR 29-JUN-1999; 99US-0140991.
```

```
PR 30-JUN-1999; 99US-0141287.
```

```
PR 01-JUL-1999; 99US-0141842.
```

```
PR 01-JUL-1999; 99US-0142154.
```

```
PR 02-JUL-1999; 99US-0142055.
```

```
PR 06-JUL-1999; 99US-0142390.
```

```
PR 08-JUL-1999; 99US-0142803.
```

```
PR 09-JUL-1999; 99US-0142920.
```

```
PR 12-JUL-1999; 99US-0142977.
```

```
PR 13-JUL-1999; 99US-0143542.
```

```
PR 14-JUL-1999; 99US-0143624.
```

```
PR 15-JUL-1999; 99US-0144005.
```

```
PR 16-JUL-1999; 99US-0144085.
```

```
PR 16-JUL-1999; 99US-0144086.
```

```
PR 19-JUL-1999; 99US-0144325.
```

```
PR 19-JUL-1999; 99US-0144331.
```

```
PR 19-JUL-1999; 99US-0144332.
```

```
PR 19-JUL-1999; 99US-0144333.
```

```
PR 19-JUL-1999; 99US-0144334.
```

```
PR 19-JUL-1999; 99US-0144335.
```

```
PR 20-JUL-1999; 99US-0144352.
```

```
PR 20-JUL-1999; 99US-0144632.
```

```
PR 20-JUL-1999; 99US-0144884.
```

```
PR 21-JUL-1999; 99US-0144814.
```

```
PR 21-JUL-1999; 99US-0145086.
```

```
PR 21-JUL-1999; 99US-0145088.
```

```
PR 22-JUL-1999; 99US-0145085.
```

```
PR 22-JUL-1999; 99US-0145087.
```

```
PR 22-JUL-1999; 99US-0145089.
```

```
PR 22-JUL-1999; 99US-0145192;
PR 23-JUL-1999; 99US-0145145;
PR 23-JUL-1999; 99US-0145218;
PR 23-JUL-1999; 99US-0145224;
PR 26-JUL-1999; 99US-0145276;
PR 27-JUL-1999; 99US-0145913;
PR 27-JUL-1999; 99US-0145918;
PR 27-JUL-1999; 99US-0145919;
PR 28-JUL-1999; 99US-0145951;
PR 02-AUG-1999; 99US-0146386;
PR 02-AUG-1999; 99US-0146388;
PR 02-AUG-1999; 99US-0146389;
PR 03-AUG-1999; 99US-0147038;
PR 04-AUG-1999; 99US-0147204;
PR 04-AUG-1999; 99US-0147302;
PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149929;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149920;
PR 23-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 18-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;

PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161992;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

Query Match 9.5%; Score 172.5; DB 21; Length 277;
Best Local Similarity 26.6%; Pred. No. 2.5e-06;
Matches 54; Conservative 34; Mismatches 74; Indels 41; Gaps 9;

QY 24 NVSAKDFGAILTYLLEEVYQGFKEARMKMAA--ASGTVSMFLYNGSEIADGRWVEVD 81
DB 48 SLSDKSSGSGHFQSNQFELYGKAEVQMKLVPNNSAGVTYTLK-----SPGTTWDEID 101
QY 82 IEVLGR---NPGFOSNITGKAGAKTSEKHHA VSPADDAFHTYGLEMTPNVYRWTVD 138
DB 102 FEFIGNISGHVYTLHTNYT-KGSGDKQGFILWFDPTAN--FHTYCITMNPORILFTVD 158
QY 139 GOEVRKTEGQVSNLTGQGLRF-----NIMSESAAWG---QFDESKLPLOF 185
DB 159 GIPIREFNMAD-----SRGVFPFKOPMRILYASIMEAHMATRGLKRTDWSKAPFTAY 212
QY 186 IN-----VKKVYKYPGGE 200
DB 213 YRNVVEGCVWNGKSVCPANSQ 235

Search completed: July 11, 2003, 10:52:52
Job time : 40 secs
```

